

Análisis de la acreditación de calidad en programas de ingeniería industrial y los resultados en las pruebas nacionales estandarizadas, en Colombia

Enrique J. Delahoz-Dominguez^(1,2), Sergio Guillen-Ibarra⁽³⁾ y Tomás Fontalvo-Herrera^{(4)*}

(1) Universidad Tecnológica de Bolívar, Facultad de Ingenierías, Programa de Ingeniería Industrial. Campus de Ternera, Cartagena - Colombia. (e-mail: edelahoz@utb.edu.co).

(2) Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018, Barcelona, España

(3) Universidad Tecnológica de Bolívar, Facultad de Ingenierías, Programa de Ingeniería Industrial. Campus de Ternera, Cartagena - Colombia. (e-mail: t00040948@utbvirtual.edu.co).

(4) Universidad de Cartagena, Facultad de Ciencias Económicas, Programa de Administración Industrial., Piedra de Bolívar Campus, Cartagena – Colombia. (e-mail: tfontalvoh@unicartagena.edu.co).

* Autor a quien debe ser dirigida la correspondencia.

Recibido Jun. 24, 2019; Aceptado Ago. 21, 2019; Versión final Sep. 6, 2019, Publicado Feb. 2020

Resumen

El presente artículo presenta un análisis relacional entre las acreditaciones universitarias sobre calidad en programas de ingeniería industrial y el resultado en las pruebas estandarizadas de la educación superior en Colombia, usando técnicas de aprendizaje automático. Primero, se desarrolló una revisión literaria concerniente a las acreditaciones sobre calidad, evaluación del desempeño organizacional y técnicas de aprendizaje automático. Segundo, se desarrolló un análisis de componentes principales para identificar factores de agrupación entre las universidades estudiadas. En tercer lugar, se evaluaron modelos de aprendizaje supervisado, para pronosticar la acreditación de una institución según los resultados en los exámenes estandarizados. Los resultados muestran que la componente de comunicación escrita es una variable crítica para determinar la clasificación de universidades como acreditadas. En conclusión, la evidencia empírica muestra la existencia de una relación compatible de contar con una acreditación sobre calidad y tener un buen desempeño en los resultados de las pruebas estandarizadas.

Palabras clave: acreditación; calidad; ingeniería industrial; educación superior; aprendizaje automático

Analysis of the quality accreditation in industrial engineering programs and the results in the standardized national tests, in Colombia

Abstract

This article presents a relational analysis between university quality accreditation in industrial engineering programs and the result in standardized tests of higher education in Colombia, employing machine learning techniques. First, a literature review was developed concerning quality accreditations, organizational performance evaluation and machine learning techniques. Then, a principal component analysis was developed to identify grouping factors among the universities studied. Third, supervised learning models were evaluated, to predict the accreditation of an institution based on the results of standardized exams. The results show how the written communication component is a critical variable to determine the classification of universities as accredited. In conclusion, the empirical evidence shows the existence of a relationship between having quality accreditation and having a good performance in the results of standardized tests.

Keywords: accreditation; quality; industrial engineering; higher education; machine learning

INTRODUCCIÓN

Actualmente las universidades enfrentan retos sociales, académicos y financieros que hacen su gestión una tarea compleja. Para Massy (2003), la credibilidad de las universidades ante la sociedad se está perdiendo, debido a una desconexión con sus principios misionales. Para una universidad moderna, la formación de profesionales competentes y la generación de conocimiento son sus objetivos fundamentales, sin embargo, otros objetivos de reconocimiento, tales como la aparición en rankings internacionales o el número de artículos publicados pasan a un primer lugar en marco de la gestión educativa.

En el anterior contexto, las acreditaciones de calidad, en su función de garantizar ante la sociedad la prestación de un servicio educativo pertinente y ajustado a las exigencias del mundo moderno, juegan un papel fundamental. Asociado a esto, está la masificación de la oferta universitaria por parte del sector privado, escenario en el cual el estado limita su participación como oferente activo del servicio de educación superior y por el contrario promueve el subsidio a la demanda como herramienta de crecimiento de la oferta, a través de la entrega de becas y créditos estudiantiles. En consecuencia, el papel del estado hace una transición desde un control centralizado de la educación superior a un modelo de supervisión (Alexander, 2000; Pineda and Celis, 2017). En donde la autonomía universitaria debe promover mecanismos exhaustivos de auditoría, verificación y seguimiento que permitan un mínimo de condiciones de calidad para la prestación del servicio educativo.

En Colombia, los procesos de apertura de un programa universitario y la certificación del mismo son procesos diferentes. En caso de la apertura, las universidades solicitan ante el estado un registro calificado para ofertar un nuevo programa, evaluando un pliego de condiciones mínimas de funcionamiento, en las cuales se involucra un estudio de mercado, infraestructura física, planta docente y estructura administrativa del programa. En cambio, el proceso de acreditación en calidad universitaria es un proceso voluntario, en el cual la universidad demuestra unas condiciones máximas de excelencia en los factores de Proyecto institucional, procesos académicos, investigación, impacto en la sociedad, visibilidad nacional e internacional, bienestar institucional, recursos financieros y planta docente. Por otra parte, el gobierno colombiano a partir del año 2017 solo brindará créditos educativos y becas a estudiantes matriculados en programas acreditados, lo cual genera un escenario de discusión sobre si existen diferencias en el aspecto misional primario de las universidades de formar profesionales competentes entre universidades acreditadas y no acreditadas.

En este contexto, las universidades deben afrontar las evaluaciones externas de la calidad de sus procesos y a su vez responder a sus objetivos internos de desarrollo, los cuales no siempre se encuentran alineados (Hladchenko, 2015). Es así como, se plantea el uso de los exámenes nacionales estandarizados como medio para evaluar el desempeño en la formación de ingenieros industriales y su relación con las acreditaciones de calidad universitarias. Como una primera aproximación al concepto de Calidad educativa asociada al desempeño académico, se ha seleccionado la profesión de Ingeniería Industrial como objeto de estudio, dado que es el programa de Ingeniería más estudiado en Colombia, lo cual permite tener un número significativo de estudiantes para el análisis propuesto. Por lo tanto, se evalúa por un lado la calidad intrínseca (Resultados del examen) y por otro la calidad extrínseca (Acreditación institucional).

Con el primer modelo se evalúa los procesos internos de formación académica, evidenciando el nivel de conocimiento del estudiante y con el modelo de acreditación se analiza cómo estos estándares enfocados en resultados externos inciden en el desempeño de las universidades objeto de estudio. En función de analizar la acreditación en calidad de las universidades y el desempeño en los exámenes nacionales estandarizados, en primer lugar, se plantea una revisión literaria de procesos de evaluación del desempeño organizacional en relación con acreditaciones y certificaciones de calidad. En segundo lugar, se presenta la metodología utilizada, describiendo las actividades y la articulación de técnicas de aprendizaje automático, por último, se plantean los resultados y hallazgos, además de la discusión de los mismos en función del contexto e investigaciones previas.

OTROS ANTECEDENTES

Las relaciones de asociación entre los sistemas de acreditación de la calidad y el desempeño organizacional han sido ampliamente estudiadas para el sector empresarial. La literatura sugiere que las acreditaciones de calidad mejoran el desempeño organizacional (Boon Wan et al., 2003; Fontalvo-Herrera and DeLaHoz-Domínguez, 2018; Linden et al., 2003). Desde otro punto de vista, las acreditaciones de calidad son consideradas ventajas competitivas (Molina-Azorín et al., 2015). Sin embargo, estudios previos demuestran empíricamente lo contrario, por ejemplo, en (Coelho et al., 2015) se analiza el desempeño financiero entre compañías certificadas en ISO: 9001 y no certificadas, encontrando que no existe diferencia significativa en el crecimiento de las ventas entre las empresas. Los resultados de (Karapetrovic et al., 2010) muestran una continua reducción de las utilidades en compañías certificadas en la norma ISO:9001.

Calidad en la educación superior

El concepto de calidad en la educación superior no es un concepto nuevo, desde su origen las universidades han trabajado en la búsqueda de la excelencia en sus actividades de docencia, investigación y proyección a la sociedad (Ruben, 2018). Sin embargo, en la actualidad el enfoque de calidad educativa se ha adaptado al concepto de calidad organizacional, desarrollado en el entorno industrial y de negocios. Es decir, se pasó de una aproximación de la calidad educativa soportada en un ideal no estructurado de mejoramiento, construido sobre evaluaciones subjetivas, a un enfoque de la calidad basado en los resultados, involucrando factores de eficiencia, productividad y mejoramiento, evaluados a través de mediciones objetivas en el tiempo.

En este mismo sentido, trabajos previos han desarrollado investigaciones similares, por ejemplo, en el análisis de la deserción estudiantil (Lykourantzou et al., 2009), utilizan Redes Neuronales para crear un indicador de severidad asociado al nivel de riesgo del alumno, por su parte Tan and Shao (2015) implementan algoritmos de aprendizaje automático en una muestra de sesenta mil estudiantes, encontrando el modelo de árbol de decisión como el de mejor efectividad. En otro enfoque, se utilizan técnicas de aprendizaje no supervisado para identificar patrones en temas educativos, es el caso de (Fontalvo-Herrera et al., 2018) donde relacionan las competencias profesionales del Instituto de Ingenieros Industriales (IES) con los planes de estudio universitarios, identificando perfiles de educación en Ingeniería Industrial utilizando clustering y PCA, también en la investigación de Ezen-Can et al., (2015) utilizan técnicas de clustering para analizar el contenido de los foros académicos y promover así soporte a los estudiantes. Otros estudios relacionados con el análisis clúster, encuentran la caracterización de perfiles de aprendizaje, tomando en cuenta una población estudiantil en la que se determinan estilos de aprendizaje dominantes, con el fin de adecuar los métodos de enseñanza existentes y mejorar en el rendimiento estudiantil (Durán and Costaguta, 2007), mientras que para (De-La-Hoz et al., 2019) los perfiles de aprendizaje son obtenidos por la interacción del estudiante con una plataforma virtual de soporte a la enseñanza.

En el análisis de estudios previos relacionados se evidencia la pertinencia de utilizar técnicas de aprendizaje automático para analizar fenómenos educativos. Por lo tanto, esta investigación analiza cómo un estándar de calidad que impacta en todo el sistema organizacional (la institución como un todo), incide en los resultados de desempeño de los procesos académicos, específicamente los resultados en las pruebas estandarizadas del estado colombiano (resultados externos de impacto en la sociedad). Consecuentemente, en esta investigación se plantean las siguientes preguntas problemáticas. ¿La acreditación de programas académicos de calidad incide en el desempeño de los resultados en los exámenes estandarizados?, ¿Existe diferencia entre los resultados obtenidos por las universidades acreditadas versus las no acreditadas?, ¿Cuáles módulos del examen estandarizado están asociados con los programas de universidades acreditadas?

MODELOS DE APRENDIZAJE SUPERVISADO

La regresión logística es una técnica estadística utilizada en esta investigación para predicciones binarias. Es necesario tener una variable dependiente y un grupo de variables independientes. Partiendo del supuesto de contar con una variable de respuesta ordinal con respecto a unos predictores, es decir esta solo puede tomar valores de 1 y 0. Una de las ventajas de este modelo es, priorizar el principal tema de interés de quien realiza el estudio, acercando las predicciones a lo que realmente se desea identificar, es especialmente útil en particular cuando solo hay dos posibles respuestas (cuando la variable de respuesta es dicotómica), que es el caso más común, la probabilidad de $Y = 1$ se usa como la variable de respuesta en un modelo de regresión:

$$p = P(Y = 1|X) = P(Y = 1|X_1, \dots, X_m) \quad (1)$$

Sea p en (1), el modelo de regresión logística de P , cuyo rango es R , como combinación lineal de las variables independientes X_1, \dots, X_m .

$$\text{logit}(P) = \ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \dots + \beta_m X_m \quad (2)$$

Esta expresión es equivalente a:

$$P = \frac{\exp(\alpha + \beta_1 X_1 + \dots + \beta_m X_m)}{1 + \exp(\alpha + \beta_1 X_1 + \dots + \beta_m X_m)} \quad (3)$$

Modelos de Árboles de decisión

Los modelos de árboles, representan un árbol donde las hojas son las categorías de las respuestas, por ejemplo, universidad acreditada (1), universidad no acreditada (0). Así, el árbol de decisión binario divide los datos (Nodo principal) en dos subconjuntos (Nodos hijos) calculando la variable que proporciona la mejor división según un criterio preestablecido. Recurrentemente, los dos nuevos subconjuntos se convierten entonces en nodos principales y subsecuentemente divididos en dos nodos hijos. Así, la división binaria continua hasta clasificar todas las observaciones. Una de las ventajas de este método es ser no-paramétrico, no se tienen en cuenta ajustes a una distribución establecida para las variables predictoras. Considerando la tendencia de los árboles de decisión a sobre-ajustarse, se usa el método de pruning, el cual consiste en la eliminación de los nodos que tienen poco impacto en el desempeño global de la clasificación (Breiman, 2017).

Random Forest (RF)

El modelo de RF es un método de ensamble, basado en la construcción recurrente y creciente de múltiples árboles de decisión a través del bootstrapping (Breiman, 2001). Es decir, se crean múltiples árboles, de diferentes composiciones de variables, de tal forma que cada árbol arroja un resultado independiente, para luego realizar un proceso de democracia donde se asigna una categoría según clase resultante más votada en general. La generación de respuestas separadas por cada árbol de decisión y su posterior ensamble en una predicción general produce modelos robustos, menos susceptibles a los valores extremos que un árbol de decisión simple, mejorando así la capacidad de predicción y clasificación del modelo. EL modelo RF presenta una técnica de selección de variables, de esta manera puede manejar grandes conjuntos de datos sin necesidad de utilizar procesos previos para la reducción de dimensiones. Además, el modelo permite identificar la importancia de cada variable para la correcta clasificación de las observaciones.

Medidas de Desempeño

El éxito del proceso de clasificación se da por la diferencia entre el valor predicho y el valor real. Esta relación se describe por las métricas Verdaderos Positivos (VP), Verdaderos Negativos (VN), Falsos Positivos (FP) y Falsos Negativos (FN). Las métricas utilizadas para evaluar el desempeño serán la tasa de clasificación correcta (C), Valor predictivo positivo (PPV), Valor predictivo negativo (PPN), Sensibilidad (S) y Especificidad (E) y el área bajo la curva (AUC). El área bajo la curva representa la tasa de TP y FP a varios umbrales de discriminación. Un modelo con una clasificación perfecta tendrá un AUC = 1. Por otro lado, un modelo totalmente aleatorio arrojará un valor de AUC = 0.5.

$$C = \frac{TP + TN}{n} \quad (4)$$

$$S = \frac{TP}{TP + FN} \quad (5)$$

$$E = \frac{TN}{TN + FP} \quad (6)$$

METODOLOGÍA

Esta investigación fue de corte evaluativa. En un segundo nivel se realizó un análisis relacional entre la acreditación de alta calidad y su impacto en las pruebas estandarizadas. La producción del conocimiento se generó a través de un análisis empírico de los desempeños en las pruebas Saber Pro y su dependencia de la Acreditación. Paralelamente, se generó a un análisis racional al articular diferentes modelos de aprendizaje automático para agrupar, clasificar y pronosticar las universidades objeto de esta investigación.

Esta investigación, parte de la conceptualización de los investigadores sobre la acreditación de calidad y su impacto en el desempeño académico, generando así las preguntas de investigación inherentes. Para esto se propone una metodología basada en aprendizaje automático de datos, mediante procesos de agrupación y clasificación se determinan relaciones de asociación y compatibilidad de ser una universidad acreditada y contar con buenos resultados en los exámenes. La primera fase consta de un análisis exploratorio de datos, posteriormente a través del Análisis de Componentes Principales (PCA) se genera una visualización en un espacio de dos dimensiones para la interpretación del desempeño de las universidades según su categoría de acreditación. Consecuentemente, se desarrolló la clasificación de universidades acreditadas y no acreditadas utilizando las técnicas de árbol de decisión (DT) y Random Forest (RF), para validar de manera objetiva los criterios de evaluación del examen. Uno de los factores claves de la metodología es la aleatorización del proceso, dado que en la predicción de categoría de acreditación se utilizan técnicas de

validación cruzada que permiten al modelo ser entrenado y evaluado en diferentes oportunidades por diferentes observaciones. La articulación de las actividades en la metodología es presentada en la Fig. 1.



Fig. 1: Estructura de Investigación

Datos

La base de datos utilizada se obtuvo de las estadísticas oficiales del Instituto Colombiano para el Fomento de la Educación Superior (ICFES). Las pruebas Saber Pro son el examen de estado de la calidad de la educación superior, el cual es realizado por (ICFES) con el fin de evaluar y proporcionar un reporte de desarrollo de las habilidades y conocimientos generales de estudiantes que han aprobado el 75% de los créditos de sus respectivos programas. El puntaje global de las pruebas Saber Pro se obtiene de la media simple de los resultados para cada competencia, Los módulos evaluados en el examen se muestran en la Tabla 1.

Tabla 1. Resumen de los módulos de la prueba Saber Pro.

| Nombre del modulo | Tipo | Símbolo | Resultado promedio acreditadas | Resultado promedio No acreditadas |
|--|------------|---------|--------------------------------|-----------------------------------|
| Inglés | Genérica | ING | 175 | 146 |
| Razonamiento Cuantitativo | Genérica | RC | 162 | 148 |
| Comunicación Escrita | Genérica | CE | 182 | 156 |
| Competencia ciudadana | Genérica | CC | 160 | 131 |
| Lectura crítica | Genérica | LC | 171 | 142 |
| Formulación de Proyectos de Ingeniería | Específica | FPI | 167 | 138 |
| Diseño de sistemas productivos y logísticos | Específica | DSPL | 166 | 141 |
| Pensamiento Científico, Matemática y Estadística | Específica | PCME | 164 | 138 |

La base de datos consta de doscientos cuarenta y cinco mil quinientos cuarenta y tres (245543) estudiantes que fueron examinados en el año 2017, de los cuales el 4.35% (10683) correspondía a estudiantes de ingeniería industrial. Se analizaron los 87 programas de ingeniería industrial, de los cuales el 31% (27 programas) contaban con la certificación de calidad. Paralelamente, 19 de estos programas correspondían a universidades públicas y 66 a privadas. Para el modelado se agregó la condición de acreditación de calidad de la universidad a través de una variable binaria (acreditada =1, no acreditada = 0).

RESULTADOS

Las diferencias de los resultados medios generales obtenidos entre universidades acreditadas y no acreditadas se evaluaron a través de un test de medias. El resultado arrojó un valor-p igual a 2.35×10^{-9} , indicando que existe compatibilidad entre la hipótesis de diferencia de resultados de la prueba entre instituciones acreditadas y no acreditadas. Consecuentemente, en la Fig. 2 se observa como las universidades acreditadas presentan mejores resultados para cada módulo evaluado en la prueba Saber Pro.

La técnica de Análisis Componentes Principales permitió generar la representación visual de los datos y crear nuevas variables no correlacionadas para mejorar los procesos posteriores de clasificación. Para la visualización, solo se seleccionaron las dos primeras componentes principales, dado que estas representan el 83% de la información total. En La Fig. 2, los puntos representan las universidades estudiadas y la figura representa la condición de acreditación.

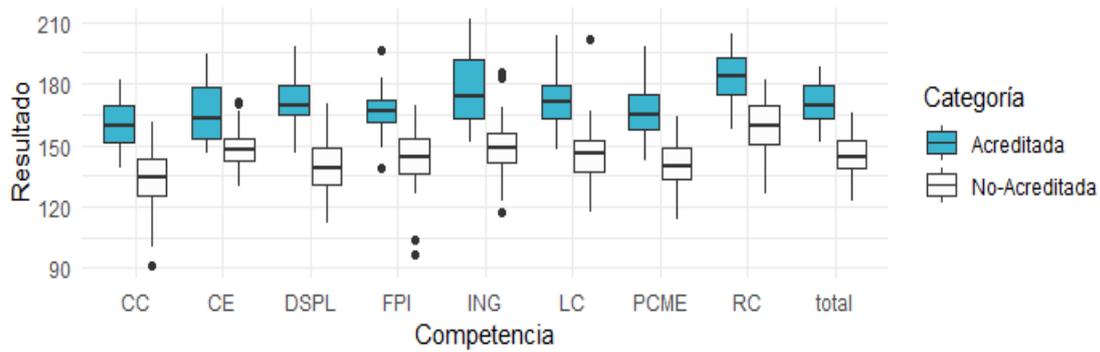


Fig. 2: Diagrama comparativo de los resultados por condición de acreditación

Así, un círculo verde representa una universidad acreditada y un triángulo naranja una No-acreditada. Igualmente, el nivel de contribución de cada variable se puede interpretar de acuerdo a la longitud de la flecha, de la forma que entre más larga sea la flecha que representa las variables, más grande será la contribución a la componente principal asociada.

En la componente 1, se observa como los módulos de competencia ciudadana, lectura crítica (LC) y pensamiento científico representan mayoritariamente la variabilidad explicada de la componente, con valores de 0,37 para cada una. Sin embargo, variables RC y DSLP e ING también cuentan con una variabilidad explicada, con valores de 0.34, 0.34 y 0.35 respectivamente. Por lo tanto, esta dimensión se interpreta como el puntaje global ponderado de la prueba. En la componente 2 se evidencia como los módulos de CE y FPI son las variables que mejor representan la variabilidad explicada de la componente con valores de -0,78 y 0,46 respectivamente, estas dos variables crean una zona de frontera sobre el cual se proyectan el resto de variables estudiadas. Además, sobre la proyección de crecimiento de CE Y FPI están contenidas el 92% de las universidades acreditadas, como se puede apreciar en la Fig. 3.

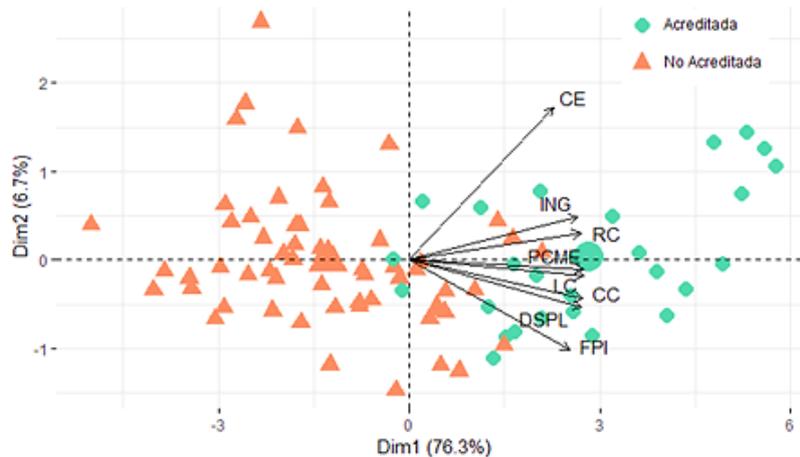


Fig. 3: Resultados de la prueba Saber Pro mediante el Análisis de Componentes Principales (PCA)

Las variables menos informativas en el PCA son (FPI), (ING) y (RC). En un análisis integral de la Fig. 3, se observa como la mayoría de las universidades no-acreditadas (triángulo) se ubica en el segundo y tercer cuadrante, esta ubicación minimiza la proyección de los módulos evaluados, dado que se encuentran en la dirección opuesta de crecimiento de la variable, asociando a esto la hipótesis que las universidades no-acreditadas tienen menores puntajes en general que las universidades acreditadas. Por su parte las universidades acreditadas se encuentran mayoritariamente en el primer y cuarto cuadrante (círculos), alineados al vector de crecimiento de cada uno de los módulos, este alineamiento permite inferir una relación compatible entre contar con una acreditación de calidad y obtener puntajes altos en los resultados del examen.

Tabla 2. Resumen de las métricas de desempeño para cada modelo

| Modelo | Min AUC | Prom AUC | Max AUC | Min S | Prom S | Max S | Min E | Prom E | Max E |
|---------------------|---------|----------|---------|-------|--------|-------|-------|--------|-------|
| Random Forest | 0.67 | 0.94 | 1 | 0.33 | 0.81 | 1 | 0.66 | 0.91 | 1 |
| Regresión Logística | 0.83 | 0.95 | 1 | 0.33 | 0.73 | 1 | 0.66 | 0.88 | 1 |
| Árbol de decisión | 0.58 | 0.83 | 1 | 0.33 | 0.75 | 1 | 0.51 | 0.9 | 1 |

Proceso de clasificación

Con el Árbol de decisión y Random Forest, se determina la categoría de acreditación de una universidad según los resultados de la prueba Saber Pro. En general, el modelo de Random Forest presenta un mejor desempeño global en la precisión del 91% y un promedio global de la métrica AUC del 94%. El resumen del desempeño de los modelos se muestra en la Tabla 2, la cual informa la variabilidad de los resultados para un mejor entendimiento del modelo. A su vez, en la Fig. 4, se observa claramente un aceptable desempeño de los modelos para clasificar las universidades según su categoría de acreditación. El modelo más robusto es el Random Forest, el cual mejora sistemáticamente a los modelos basados Árboles de decisión.

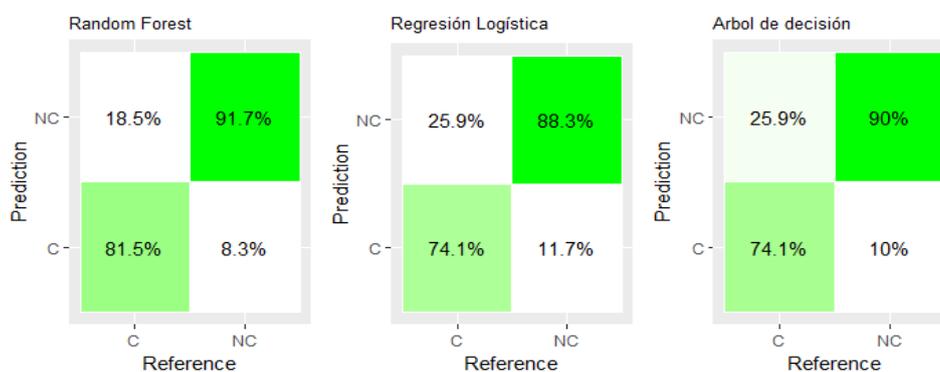


Fig. 4: Matriz de confusión para el proceso de validación.

DISCUSIÓN FINAL

El modelo de aprendizaje automático presentado, analiza la relación entre las acreditaciones de calidad universitarias y los resultados de un examen estandarizado, para el caso de los programas de ingeniería industrial en Colombia. El modelo logra altos niveles de precisión, superiores a los encontrados por Huang and Fang (2013) para la predicción del desempeño en cursos de ingeniería, utilizando como objeto de estudio a 323 estudiantes de grado, la precisión de este modelo es (81% – 91%) IC.

Como hallazgo significativo, se demostró como la comunicación escrita (CE) es la frontera que divide las universidades según su condición de acreditación, además de definir el posicionamiento entre universidades con altos y bajos resultados. Generando así, evidencia empírica para soportar las teorías de (Cilliers, 2012), donde propone el mejoramiento de las habilidades de escritura para un mejor desempeño en estudiantes de cursos de programación de computadores. Por su parte Jefferies et al., (2018) muestra como la enseñanza de habilidades de escritura puede mejorar el desempeño profesional de los estudiantes universitarios. De forma similar, otras investigaciones muestran como una buena habilidad de escritura y el mejoramiento del desempeño académico tienen una relación positiva (Shao and Purpur, 2016; Yang, 2016). En el contexto colombiano, el proceso sistemático de evaluación de la calidad universitaria arroja que para el objetivo misional de formación de profesionales competentes, las universidades acreditadas presentan mejores resultados, esto como una posible valoración positiva de la sociedad de la certificación de calidad universitaria, redundando en una mejor percepción de los estudiantes al momento de elegir una universidad y la conciencia colectiva de mejoramiento plasmada en los planes de mejoramiento para sostener la certificación de calidad.

El presente estudio sigue los lineamientos de la Asociación Estadística Americana en su estudio sobre el uso del valor-p para análisis relacionales (Wasserstein et al., 2019), analizando para este caso la relación entre las acreditaciones de calidad en las universidades y los resultados de las pruebas estandarizadas nacionales desde diferentes perspectivas, siempre buscando una aproximación al conocimiento, integrando la información aportada por los modelos de aprendizaje automático para aportar nueva información en esta área del conocimiento. Es así como el enfoque propuesto genera un contexto de análisis para relacionar positivamente el hecho de contar con una acreditación de calidad con tener buenos resultados en las pruebas estandarizadas en programas de ingeniería industrial.

CONCLUSIONES

1.- Esta investigación, evidencia empíricamente la relación existente entre acreditaciones de calidad universitarias y el desempeño académico para programas de ingeniería industrial en Colombia. Lo anterior, fue exitosamente modelado a través de la articulación de diferentes técnicas de aprendizaje automático. De igual manera, los resultados mostraron como hallazgo importante, que el desempeño en la evaluación de Comunicación escrita incide de manera significativa en que las universidades estén acreditadas o no.

2.- Se aporta conocimiento al concepto de calidad educativa, ratificando que los procesos administrativos y operativos propuestos por un modelo gestión de alta calidad tienen impacto directo en el desempeño de las pruebas de estado. Incidiendo en la formación de profesionales con los conocimientos necesarios para poder desempeñarse satisfactoriamente en su contexto profesional y laboral.

3.- Se aporta un modelo robusto de pronóstico, con un alto porcentaje medio de precisión para determinar la acreditación de calidad de una universidad usando los resultados de una prueba estandarizada. Factor clave para pronosticar escenarios de desarrollo y planeación para los tomadores de decisión en las universidades.

REFERENCIAS

- Alexander, F. K. The Changing Face of Accountability. *The Journal of Higher Education*, 71(4), 411–431 (2000)
- Boon Wan, T., M. Goh y C. Chow-Chua. Does ISO 9000 certification improve business performance? *International Journal of Quality & Reliability Management*, 20(8), 936–953 (2003)
- Breiman, L. Random Forests. *Machine Learning*, 45(1), 5–32 (2001)
- Breiman, L. *Classification and Regression Trees*. (2017)
- Cilliers, C. B. (2012). Student perception of academic writing skills activities in a traditional programming course. *Computers & Education*, 58(4), 1028–1041 (2012)
- Coelho, L. M. S., C.J.F Cândido y R.MT. Peixinho, The financial impact of a withdrawn ISO 9001 certificate. *International Journal of Operations & Production Management*, 36(1), 23–41. <https://doi.org/10.1108/IJOPM-11-2014-0540> (2015)
- De-La-Hoz, E., E. De-La-Hoz y T. Fontalvo. Metodología de Aprendizaje Automático para la Clasificación y Predicción de Usuarios en Ambientes Virtuales de Educación. *Inf. Tecnológica* 30, 247–254 (2019)
- Durán, E., y R. Costaguta, Minería de datos para descubrir estilos de aprendizaje. *Revista Iberoamericana de Educación*, 42(2), 6 (2007)
- Ezen-Can, A., K.E Boyer, S. Kellogg y S. Booth, Unsupervised modeling for understanding MOOC discussion forums: a learning analytics approach. *Proceedings of the fifth international conference on learning analytics and knowledge*, 146–150 (2015)
- Fontalvo-Herrera, Tomás J., E.J Delahoz y A.A Mendoza-Mendoza, Aplicación de Minería de Datos para la Clasificación de Programas Universitarios de Ingeniería Industrial Acreditados en Alta Calidad en Colombia. *Información tecnológica*, 29(3), 89–96 (2018)
- Fontalvo-Herrera, Tomás José, y E.J DeLaHoz-Domínguez, Study of financial efficiency in companies certified with the BASC label using Data Envelopment Analysis: Case applied in Cali - Colombia. *Entramado*, 14(1), 78–87 (2018)
- Hladchenko, M, Balanced Scorecard – a strategic management system of the higher education institution. *International Journal of Educational Management*, 29(2), 167–176 (2015)
- Huang, S y N. Fang, Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 61, 133–145 (2013)
- Jefferies, D., S. McNally, y otros cinco autores, The importance of academic literacy for undergraduate nursing students and its relationship to future professional clinical practice: A systematic review. *Nurse Education Today*, 60, 84–91 (2018)
- Karapetrovic, S., M.C Fa y I.H Saizarbitoria, What happened to the ISO 9000 lustre? An eight-year study. *Total Quality Management & Business Excellence*, 21(3), 245–267 (2010)
- Linden, G., B. Smith y J. York, Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, 7(1), 76–80 (2003)
- Lykourantzou, I., Giannoukos y otros tres autores, Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education*, 53(3), 950–965 (2009)
- Massy, W. F. Honoring the trust: Quality and cost containment in higher education. ERIC (2003)
- Molina-Azorín, J. F., J.J. Tarí y otros tres autores, The effects of quality and environmental management on competitive advantage: A mixed methods study in the hotel industry. *Tourism Management*, 50, 41–54 (2015)
- Pineda, P y J. Celis, ¿Hacia la universidad corporativa? Reformas basadas en el mercado e isomorfismo institucional en Colombia. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, (25), 1–32 (2017)
- Ruben, B.D. *Quality in higher education*. Routledge. (2018)
- Shao, X. y G. Purpur, G, Effects of Information Literacy Skills on Student Writing and Course Performance. *The Journal of Academic Librarianship*, 42(6), 670–678 (2016)
- Tan, M. y P. Shao, Prediction of student dropout in e-learning program through the use of machine learning method. *International Journal of Emerging Technologies in Learning (iJET)*, 10(1), 11–17 (2015)
- Wasserstein, R. L., A.L. Schirm y N.A. Lazar, Moving to a World Beyond “p < 0.05”. *The American Statistician*, 73(sup1), 1–19 (2019)
- Yang, Y.-F, Transforming and constructing academic knowledge through online peer feedback in summary writing. *Computer Assisted Language Learning*, 29(4), 683–702 (2016)