



Enfoque de minería de datos para predecir el desempeño de estudiantes de primer año en una universidad usando los requisitos de admisión

Aderibigbe Israel Adekitan¹ y Etinosa Noma-Osaghae¹

Recibido: 17 de septiembre de 2018 / Aceptado: 13 de noviembre de 2018 / Published online: 03 December 2018
Springer Science+Business Media, LLC, parte de Springer Nature 2018

Resumen

El rendimiento académico de un estudiante en una universidad está determinado por una serie de factores, tanto académicos como no académicos. Los estudiantes que anteriormente se destacaron en el nivel de la escuela secundaria pueden perder el enfoque debido a la presión de los compañeros y el estilo de vida social, mientras que aquellos que anteriormente tuvieron problemas debido a las distracciones familiares pueden concentrarse fuera de casa y, como resultado, sobresalir en la universidad. La admisión a la universidad en Nigeria se basa típicamente en las características cognitivas de entrada de un estudiante, que es en su mayoría académica, y no necesariamente se traduce en excelencia una vez en la universidad. En este estudio, se examinó la relación entre los requisitos de ingreso de admisión cognitiva y el rendimiento académico de los estudiantes en su primer año, usando su CGPA y clase de grado usando seis algoritmos de minería de datos en las plataformas KNIME y Orange. Se observaron precisiones máximas de 50,23 % y 51,9 % respectivamente, y los resultados se verificaron mediante modelos de regresión, con valores de R^2 de 0,207 y 0,232 registrados que indican que el desempeño de los estudiantes en su primer año no se explica completamente por los requisitos cognitivos de ingreso.

Palabras clave Rendimiento académico. Aprendizaje automático. Minería de datos educativos. Algoritmos de minería de datos. Descubrimiento del conocimiento. universidad nigeriana

1. Introducción

Nigeria, con una población repleta de más de 180 millones de personas, tiene un grupo muy grande de personas que desean sinceramente una educación universitaria. La última década ha sido testigo de un fuerte aumento en la demanda de educación universitaria en Nigeria (Aina 2002). La demanda es tan

* Aderibigbe Israel Adekitan
aderibigbe.adekitan@covenantuniversity.edu.ng

¹ Departamento de Ingeniería Eléctrica y de la Información, Universidad Covenant, Ota, Ogun Estado, Nigeria

tan alto que solo un exiguo 15,3% de los solicitantes como máximo tienen la oportunidad de ser admitidos en una universidad nigeriana cada año (Aluede et al. 2012). Nigeria cuenta actualmente con ciento cincuenta y dos (152) universidades que satisfacen las demandas de educación superior de 1,7 millones de aspirantes a estudiantes nuevos cada año. Esto es además de los quinientos mil (como mínimo) jóvenes nigerianos que se gradúan de estas instituciones superiores de aprendizaje cada año (Saint et al. 2003). Las universidades nigerianas tienen personas de diversas culturas que se unen para aprender. La región del norte de Nigeria, que antes estaba relajada con respecto a la educación universitaria, ahora está presenciando un aumento en la demanda de educación universitaria (Popoola et al. 2018).

Con el crecimiento creciente de la población nacional y la demanda de educación de posgrado, se hizo evidente que el número de instituciones terciarias en el país es sumamente inadecuado. El Gobierno Federal de Nigeria ha explorado el aprendizaje electrónico y a distancia para atender a la gran cantidad de nigerianos que desean una educación universitaria. La Universidad Nacional Abierta de Nigeria (NOUN) se estableció principalmente con el fin de brindar servicios de educación a distancia y aprendizaje electrónico a los nigerianos y otras partes interesadas en el extranjero (Ajadi et al. 2008). NOUN sirve como una plataforma de aprendizaje abierto y a distancia (ODL) que proporciona materiales de instrucción a los estudiantes, para permitir el aprendizaje a su propio ritmo, lo cual es muy conveniente para los estudiantes que trabajan según el marco de la Universidad Nacional Abierta Indira Ghandi (IGNOU), India. El programa necesita un monitoreo y una mejora continuos del desempeño para garantizar el cumplimiento de las mejores prácticas globales en la entrega de ODL. La implementación de sesiones de estudio asistidas por pares en línea (Nikolic y Nicholls 2018) podría ayudar a proporcionar instrucción complementaria y apoyo entre pares para los estudiantes de NOUN para mejorar el rendimiento de los estudiantes.

Según (Olsson y Mozelius 2016), cuando la facilitación directa del aprendizaje no es factible, es necesario garantizar que las plataformas alternativas de autoaprendizaje sean adecuadas para evitar que aumenten las tasas de fracaso. El estudio de (Burke y Fedorek 2017; Kurt 2017) afirma que el aprendizaje autodirigido, basado en el modelo de aula invertida, mejora el aprendizaje de los estudiantes, mientras que (Van WYK 2018) recomienda el uso del concepto de aula invertida en programas abiertos de aprendizaje electrónico a distancia para mejorar a los estudiantes. ' experiencia, comprensión y rendimiento. Por el contrario, (Burke y Fedorek 2017; Ebbeler 2013) postulan que el aula invertida puede no mejorar necesariamente la experiencia de aprendizaje de los estudiantes porque los estudiantes que están acostumbrados a las clases tradicionales a menudo prefieren el enfoque tradicional, mientras que los que están en clases invertidas pueden no estar preparados para la transición. Además, puede ser un desafío enseñar una clase invertida y, del mismo modo, las clases invertidas pueden no ser fáciles para los estudiantes.

La desregulación de la educación universitaria en Nigeria trajo un nuevo impulso para la creación de más universidades en Nigeria (Adeogun et al. 2009). Pero este impulso vino con los problemas concomitantes de una disparidad más clara entre los que tienen y los que no, y la calidad educativa en declive debido a la maximización de las ganancias (Babalola 1998). El Gobierno Federal de Nigeria, a través del instrumento de la Comisión Nacional de Universidades (NUC), sigue trabajando arduamente para aliviar la miríada de problemas que aquejan a la educación universitaria en Nigeria. El organismo proporciona las necesidades regulatorias para la educación universitaria en Nigeria y garantiza, junto con otras partes interesadas, la implementación de todas las políticas de garantía de calidad en las universidades nigerianas (Ajayi y Ekundayo 2008). La Comisión Nacional de Universidades junto con otros organismos como la Junta Conjunta de Admisiones y Matriculación (JAMB), el Comité Nacional de Exámenes

(NECO) y el Consejo de Exámenes de África Occidental (WAEC) trabajan de la mano para establecer puntos de referencia y estándares para la admisión en las universidades de Nigeria. Por lo general, se utilizan muchos criterios, siendo el más destacado la necesidad de tener créditos en al menos cinco (5) materias, Matemáticas e Inglés inclusive, en un examen WAEC o NECO reciente (Adeyemi 2001).

Recientemente, muchas universidades iniciaron programas internos de selección universitaria de pregrado en función de sus criterios personalizados. Un ejemplo famoso es el programa de Evaluación de Aptitud Académica de la Universidad de Covenant (CUSAS) (Popoola et al. 2018). El CUSAS de la Universidad Covenant, por ejemplo, tiene una lista muy estricta de requisitos para sus aspirantes a estudiantes universitarios. Estos requisitos abarcan áreas académicas, emocionales, financieras, sociales, morales y espirituales (Popoola et al. 2018). El objetivo de los programas y políticas de selección para la admisión de estudiantes es garantizar que los estudiantes completos, en términos de capacidad, capacidad, creatividad y motivación, obtengan las pocas plazas de admisión disponibles, prestigiosas, entusiastas y justas.

En este estudio, la relación entre las características cognitivas de ingreso de los estudiantes en el momento de la admisión, medidas por la edad de ingreso de los estudiantes, el puntaje WAEC agregado, el puntaje JAMB, el puntaje CUSAS basado en la universidad y el rendimiento académico del primer año de los estudiantes medido por su clase de grado y el CGPA real se consideran utilizando la minería de datos. Este estudio busca determinar el alcance de la relación entre los criterios de admisión utilizados por las universidades nigerianas para seleccionar a los posibles solicitantes de pregrado calificados para la admisión de 100 L y el rendimiento académico del estudiante después de la primera sesión académica utilizando algoritmos de minería de datos y análisis basados en regresión. modelos

2 Minería de datos educativos

La minería de datos es un proceso de descubrimiento de conocimiento que implica la extracción de información inteligente de un conjunto de datos determinado utilizando metodologías científicas (Azevedo 2018; Hussain et al. 2018). El conjunto de datos acumulado en el tiempo extra de un proceso o sistema contiene información histórica oculta que se puede extraer para mejorar la calidad de los procesos de toma de decisiones. La minería de datos implica el uso de algoritmos para identificar patrones y tendencias dentro de un conjunto de datos. La minería de datos educativos es la extracción de información útil de un conjunto de datos generado en el dominio educativo (Tair y El-Halees 2012; Ryan y Baker 2010; Senthil y Lin 2018; Bharara et al. 2018). Las bases de datos educativas son ricas fuentes de información para evaluar el desempeño de los estudiantes y para varios análisis predictivos (Ahmed y Elaraby 2014; Khedr y El Seddawy 2015; Bharara et al. 2018). También ayuda a identificar cualquier relación oculta entre el desempeño de los estudiantes y sus características de aprendizaje (Hussain et al. 2019; Ahuja et al. 2019) y comportamientos (Kim et al. 2018). La información útil obtenida de la minería de datos se puede utilizar para mejorar la calidad y el modo de impartir los sistemas de educación superior (Daradoumis et al. 2019) con el fin de mejorar la eficiencia de la enseñanza y, en última instancia, el rendimiento de los estudiantes (Baeppler y Murdoch 2010; Osmanbegović y Suljic 2012; Agarwal et al. 2012). La información extraída de datos puede servir de guía para identificar áreas de intervención, como el rediseño de cursos, la modificación de estrategias de comunicación, la mejora de los métodos de evaluación, etc., para mejorar la calidad de la educación y la aptitud (Baeppler y Murdoch 2010).

La minería de datos educativos es un proceso basado en datos para identificar los problemas de aprendizaje de los estudiantes y las tendencias de desempeño en las instituciones de aprendizaje (Bucos y Drȃgulescu 2018). El aprendizaje automático ha encontrado aplicación en el estudio del comportamiento académico y la predicción del rendimiento de los estudiantes (Kostopoulos et al. 2018). Las diversas áreas de aplicación del aprendizaje automático en la educación incluyen la predicción del potencial de abandono y graduación (Ahuja y Kankane 2018; Nurhayati et al. 2018), la predicción del rendimiento académico (Roy y Garg 2018; Fernandes et al. 2019), la evaluación de el proceso de aprendizaje (Khan y Ghosh 2018) e identificación de riesgos de aprendizaje, evaluación de los administradores, evaluación de la retroalimentación textual de los estudiantes (Ibrahim et al. 2019; Atta UR et al. 2018), evaluación de las interacciones entre los actores educativos, como un piloto para guiar la implementación e integración de tecnologías educativas basadas en instituciones (Angeli et al. 2017), y así sucesivamente (Rodrigues et al. 2018). El aprendizaje basado en datos está a la orden del día, y se están dedicando más trabajos de investigación a mejorar el conocimiento en este campo. En minería de datos educativos, la precisión del algoritmo y el modelo implementado dependen de la naturaleza y el tamaño de los datos en estudio (Almarabeh 2017). La minería de datos educativos y la minería de datos en general están fuertemente impulsadas por los datos y, como tales, pueden surgir problemas de privacidad y consentimiento con respecto al valor comercial potencial de los datos (Lynch 2017). Por lo tanto, deben garantizarse modalidades adecuadas para su gestión.

3 Estudios relacionados

El estudio de (Baeppler y Murdoch 2010) identificó una distinción entre el análisis académico y la minería de datos. El análisis académico es un proceso basado en hipótesis que implica el uso de un conjunto de datos para resolver un problema académico, mientras que la minería de datos es un proceso especulativo para identificar una gran cantidad de información útil a partir de información aparentemente insensible sin ninguna hipótesis predeterminada.

Según (Bharara et al. 2018), el análisis de aprendizaje implica la aplicación de herramientas analíticas para estudiar a los estudiantes, la recopilación y el análisis de datos relacionados con la interacción de la plataforma de aprendizaje virtual del estudiante, mientras que la minería de datos educativos es un enfoque de aprendizaje basado en computadora para detectar únicos. patrones en el conjunto de datos relacionado con el estudiante para identificar nuevos hallazgos y para probar teorías.

(Ahmed y Elaraby, 2014) realizaron un análisis predictivo utilizando 1548 registros de muestra de estudiantes de 2005 a 2010. En el estudio, se aplicaron características como la calificación de la prueba de laboratorio, la calificación parcial, el nivel de participación de los estudiantes, etc., utilizando un árbol de decisión para identificar a los estudiantes que tienen probabilidad de reprobación. (Bhise et al. 2013) aplicó el algoritmo de agrupación K-means para predecir el resultado de los estudiantes utilizando los registros de la base de datos de los estudiantes, mientras que en el estudio de (Yadav et al. 2012), el potencial de abandono de un estudiante se predijo utilizando un árbol de decisión . clasificadores. Usando características seleccionadas de la base de datos de los estudiantes, el desempeño de los estudiantes fue clasificado por (Kaur y Kaur 2018) usando seis algoritmos de minería de datos diferentes; Naive Bayes, K-NN, 1-NN y árbol de decisión (C5.0, C4.5, CART). Mediante la aplicación de técnicas de descubrimiento de conocimiento mediante minería de datos, el estudio de (Burgos et al. 2018) evaluó registros de estudiantes de plataformas de aprendizaje electrónico para estudiantes que toman cursos de educación a distancia utilizando modelos predictivos. Después de la implementación del modelo, se desarrolló e implementó un plan de enseñanza, y el plan de enseñanza pudo reducir la tasa de deserción de los estudiantes en un 14% en comparación con las sesiones académicas anteriores.

En el estudio de (Bucos y Drăgulescu 2018), la extracción de datos educativos se llevó a cabo utilizando 908 muestras de datos de estudiantes preprocesados y cinco algoritmos de clasificación. Las características consideradas incluyen la información del curso, la asistencia de los estudiantes, las calificaciones promedio, el nivel de actividad de los estudiantes y el número de créditos aprobados en la sesión anterior. (Pelumi Oguntunde et al. 2018) [demostró el uso de correlación y análisis de regresión lineal múltiple para predecir el CGPA de graduación de los estudiantes utilizando su GPA durante un período de cinco años](#). En (Sivakumar y Selvaraj 2018), el rendimiento de los estudiantes se predijo utilizando múltiples clasificadores supervisados para comparar el rendimiento. En el estudio, se identificaron cuatro categorías de desempeño que van desde excelente a pobre. Utilizando 300 registros de muestra de estudiantes de 3 universidades, (Hussain et al. 2018) llevó a cabo un análisis predictivo del rendimiento de los estudiantes en la plataforma WEKA mediante el análisis de 12 características significativas utilizando 4 algoritmos de clasificación.

4 Metodología

En este estudio, se llevó a cabo un análisis para mostrar el alcance de la relación entre el expediente académico de los estudiantes de la Universidad Covenant en Nigeria en el momento de la admisión, en función de los requisitos de ingreso a la universidad, y el rendimiento académico de los estudiantes admitidos en el primer año medido por su CGPA de 100 L y la clase de grado utilizando modelos de regresión y minería de datos predictivos. En esta área se han implementado métodos como el análisis de regresión y las redes neuronales artificiales (Arsad y Buniyamin 2014), pero en este estudio, se implementó un modelo de minería de datos en KNIME y la plataforma Orange, y se verificó la veracidad de la predicción mediante el análisis de regresión para comparación de rendimiento. Según el conjunto de datos de (Odukoya et al. 2018), se analizaron 1445 registros de estudiantes de 2005 a 2009. Se examinaron las siguientes características en el conjunto de datos: la edad de ingreso del estudiante, el puntaje WAEC agregado, el puntaje JAMB, el puntaje CUSAS basado en la universidad, la clasificación de grado del primer año, mientras que el CGPA real se consideró para el análisis de regresión. Estas características representan los requisitos clave de la Comisión Universitaria de Nigeria (NUC) para la admisión a programas de ingeniería en Nigeria. Las calificaciones de los estudiantes se dividen en las siguientes categorías, primera clase (1ra), segunda clase inferior (2|1), segunda clase superior (2|2) y tercera clase (3ra) respectivamente. Las aplicaciones KNIME y Orange Analytic se ejecutaron en una computadora Core™ i3-7100 U CPU 2.4GHz, 6 GB Ram con sistema operativo Windows 10.

5 Estadísticas descriptivas del conjunto de datos

Esta sección presenta las estadísticas descriptivas del conjunto de datos de estudiantes analizados en este estudio. La Tabla 1 muestra las estadísticas descriptivas de las variables numéricas en el conjunto de datos. En las Figs. 1, 2, 3 y 4 se intentó ajustar los datos utilizando 5 distribuciones diferentes. Las Figuras 1 (a) y 1 (b) presentan la probabilidad acumulada y la función de densidad de probabilidad para los datos de CGPA, mientras que las Figuras 2 (a) y (b) muestran la probabilidad acumulada y la función de densidad de probabilidad para los datos de puntuación CUSAS para todos los 1445 estudiantes. En las Figs. 3 (a) y (b), la probabilidad acumulada y los gráficos de la función de densidad de probabilidad para la puntuación JAMB son

Tabla 1 Estadísticas descriptivas de los requisitos de ingreso

	Min.	Máx.	Media	estándar desviación	Diferencia	Oblicuidad	curtosis
Edad de entrada	15	24	17.9232	1.2067	1.4560	0.9648	1.5256
Puntuación JAMB	133	298	222.7723	23.2200	539.1690	$\bar{y}0.0798$	$\bar{y}0.1973$
Puntaje CUSAS	41	110	69.4999	8.9183	79.5363	$\bar{y}0.0232$	0.0136
CGPA	1.8	4.93	3.5853	0.6394	0.4088	$\bar{y}0.2462$	$\bar{y}0.4884$
WAEC	1.67	4.88	3.2566	0.6068	0.3682	0.0508	$\bar{y}0.6116$

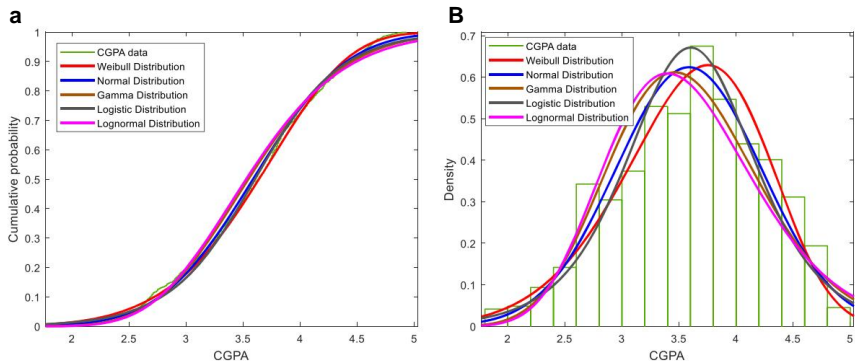


Fig. 1 Gráfico de datos de CGPA que muestra (a) Probabilidad acumulada (b) Función de densidad de probabilidad

presentado mientras que las Figs. 4 (a) y 4 (b) muestran la probabilidad acumulada y la gráfica de densidad de probabilidad de la puntuación WAEC agregada. Las características de los datos analizados se presentan como diagramas de caja que clasifican las puntuaciones de los estudiantes en función de la clase

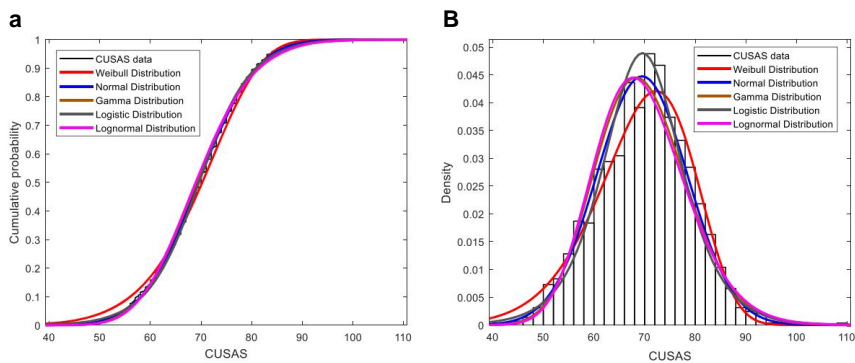


Fig. 2 Gráfico de datos CUSAS que muestra la (a) Probabilidad acumulada (b) Función de densidad de probabilidad

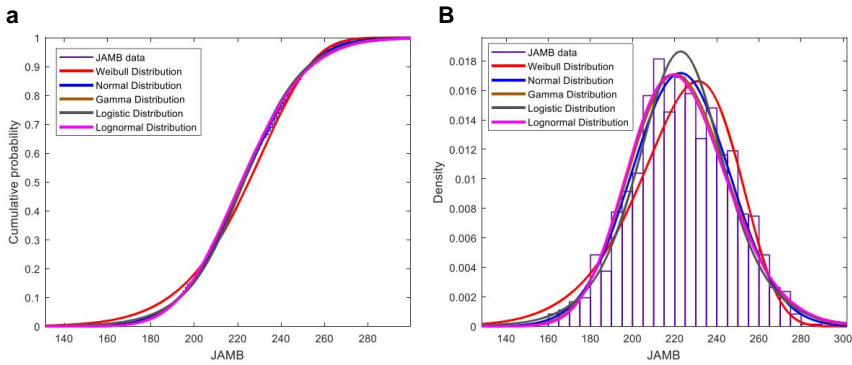


Fig. 3 Diagrama de datos de JAMB que muestra (a) Probabilidad acumulada (b) Función de densidad de probabilidad

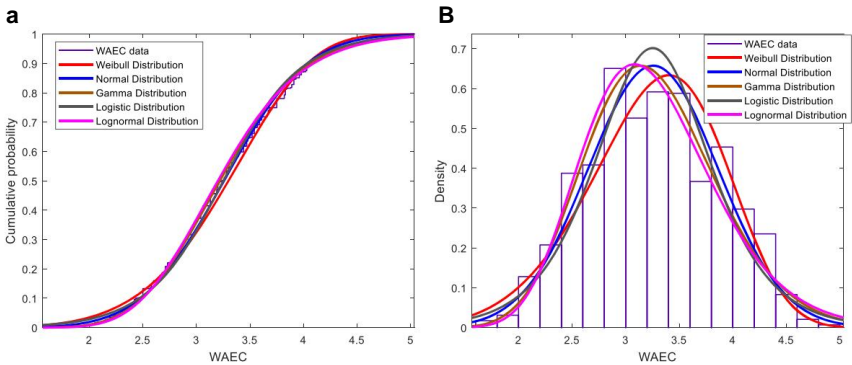


Fig. 4 Gráfico de datos WAEC que muestra la (a) probabilidad acumulada (b) función de densidad de probabilidad

de grado (1º, 2º y 3º curso) del alumno para el primer año. La figura 5 (a) presenta el diagrama de caja de la puntuación CUSAS, la figura 5 (b) presenta la puntuación JAMB,

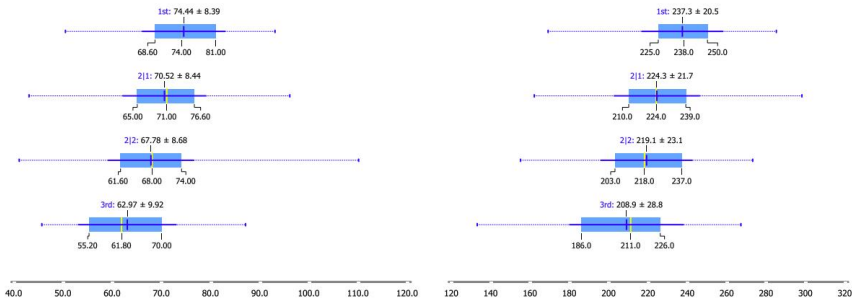


Fig. 5 Diagrama de caja que muestra la clasificación de datos por clase de grado para (a) puntaje CUSAS (b) puntaje JAMB

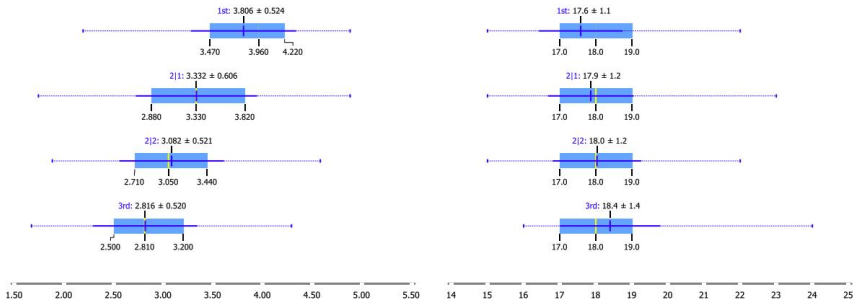


Fig. 6 Diagrama de caja que muestra la clasificación de datos por clase de grado para (a) puntaje WAEC (b) puntaje AGE

La Fig. 6 (a) muestra el diagrama de caja de la puntuación WAEC agregada y la Fig. 6 (b) muestra el diagrama de caja de la función de datos AGE.

6 Los resultados del análisis predictivo usando la plataforma KNIME

El modelo predictivo basado en KNIME implementado en este estudio se muestra en la Fig. 7. Las muestras se seleccionaron mediante muestreo estratificado y el 70 % de la muestra de datos se utilizó para entrenar el modelo mediante 6 algoritmos de extracción de datos; los algoritmos Random Forest, Tree Ensemble, Decision Tree, Naive Bayes, Logistic Regression y Resilient back propagation (Rprop) Multi-Layer Perceptron. El 30% restante de las muestras se utilizó para evaluar el rendimiento del modelo. Se llevó a cabo una reducción de dimensiones mediante el análisis de componentes principales para mejorar la precisión del modelo.

Las siguientes características de requisitos de entrada de admisión se aplicaron en el modelo de minería de datos; la edad de ingreso del estudiante, el puntaje WAEC agregado, el puntaje JAMB y el puntaje CUSAS basado en la universidad para predecir la clase del primer año del estudiante. En la Tabla 2 se presenta un análisis comparativo del rendimiento de los seis modelos. El algoritmo de regresión logística tuvo la precisión de predicción más alta con un 50,23 %, seguido

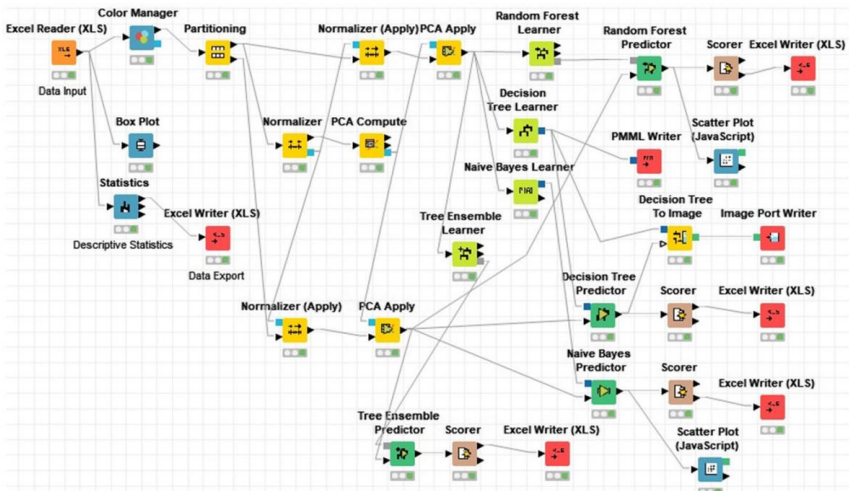


Fig. 7 El flujo de trabajo de KNIME

Tabla 2 Comparación del rendimiento del modelo

	Aleatorio bosque	Árbol Juntos	Decisión Árbol	Ingenuo Bayés	Logístico Regresión	Rprop. MLP
Correcto clasificado	210	205	172	186	218	213
Exactitud	48,38%	47,24%	39.631%	42.857%	50,23%	49.078%
Kappa de Cohen (k)	0.09	0.071	0.0041	0.106	0.081	0.084
mal clasificado	224	229	262	248	216	221
Error	51.613%	52.765%	60.368%	57.143%	49,77%	50,922%

por el algoritmo Rprop MLP con una precisión de 49.078% y esto es seguido por el Algoritmo Random Forest con 48.38% de precisión. El siguiente en jerarquía es el Árbol Algoritmo de conjunto con una precisión del 47,24%, y esto es seguido por el Naive Algoritmo de Bayes con una precisión del 42,857%. El algoritmo del árbol de decisiones tenía la precisión mínima de 39.631%. La Tabla 2 muestra comparativamente las variaciones en el número de muestras correctamente y mal clasificadas por los seis algoritmos.

La precisión de los modelos es mucho menor de lo esperado, y esta expectativa se debe a que de la suposición común de que el desempeño académico de un estudiante basado en el ingreso requisitos es un fuerte indicador del desempeño del estudiante una vez en la universidad sistema. Aunque, esta creencia puede no estar totalmente equivocada, pero el alcance de su validez necesita estar determinado. Por ejemplo, en las Figs. 5 (a), 5 (b) y 6 (a) usando la clase del primer año grados (1°, 2°|1, 2°2 y 3° grado) se puede ver en los diagramas de caja que el orden de los puntajes más altos en cada grupo sigue el orden de la clase de calificaciones para el CUSAS, el Las puntuaciones de JAMB y WAEC respectivamente, lo que implica que, de hecho, existe una relación entre la excelencia académica basada en los requisitos de admisión y el desempeño después. La Figura 6 (b) muestra que no hay una variación significativa en las edades de los estudiantes para toda la clase de grados. La Tabla 3 compara el Verdadero Positivo y el Falso Positivo predicciones para los 6 algoritmos de minería de datos considerados en el modelo KNIME.

7 Los resultados del análisis predictivo utilizando la plataforma Orange

Para verificar la veracidad del desempeño del modelo KNIME se realizó un análisis similar realizado utilizando la plataforma de minería de datos de Orange como se muestra en la Fig. 8. Seis minería de datos

Tabla 3 Confusión de predicción de los tres predictores de minería de datos

Árbol de bosque aleatorio Árbol de decisión de conjunto Regresión logística bayesiana ingenua Rprop. MLP												
	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP
3ro	0	5	0	6	2	17	8	61	0	0	0	1
2 1	134	128	127	125	101	122	125	108	164	152	140	129
2 2	73	80	76	90	59	92	41	54	53	63	72	88
1º	3	11	2	8	10	31	12	25	1	1	1	3
Total	210	224	205	229	172	262	186	248	218	216	213	221

TP Verdadero Positivo, FP Falso Positivo

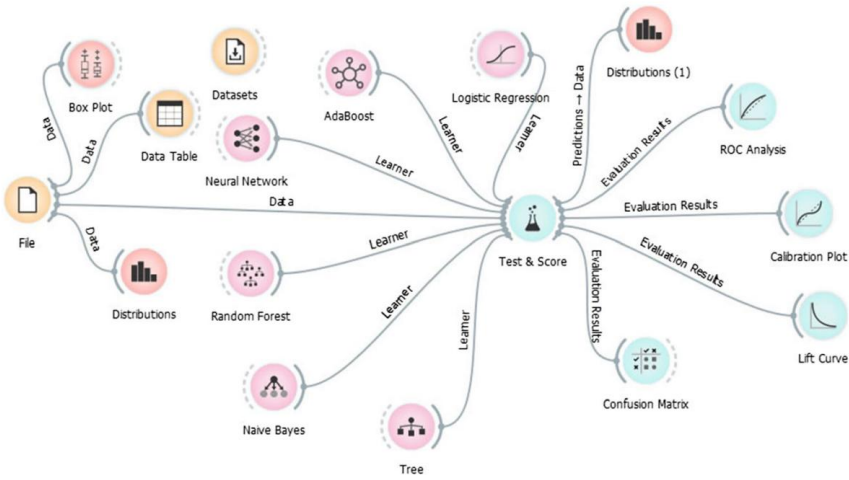


Fig. 8 El flujo de trabajo del modelo Orange

Se consideraron algoritmos para el modelo de aprendizaje supervisado, y estos son: el Árbol, el Random Forest, la Red Neuronal, el Naïve Bayes, la Regresión Logística y el algoritmo AdaBoost respectivamente. El rendimiento de cada uno de los algoritmos es presentado comparativamente en la Tabla 4. La Red Neuronal tuvo la clasificación más alta

Tabla 4 Comparación de rendimiento de los algoritmos de minería de datos en la plataforma Orange

Método	ABC	(ESE)	F1	Precisión	Recuerdo
Árbol	0.552	0.436	0.436	0.439	0.436
Bosque aleatorio	0.607	0.503	0.484	0.489	0.503
Red neuronal	0.635	0.519	0.494	0.486	0.519
bayesiana ingenua	0.642	0.500	0.489	0.486	0.500
Regresión logística	0.644	0.511	0.463	0.52	0.511
AdaBoost	0.531	0.428	0.429	0.430	0.428

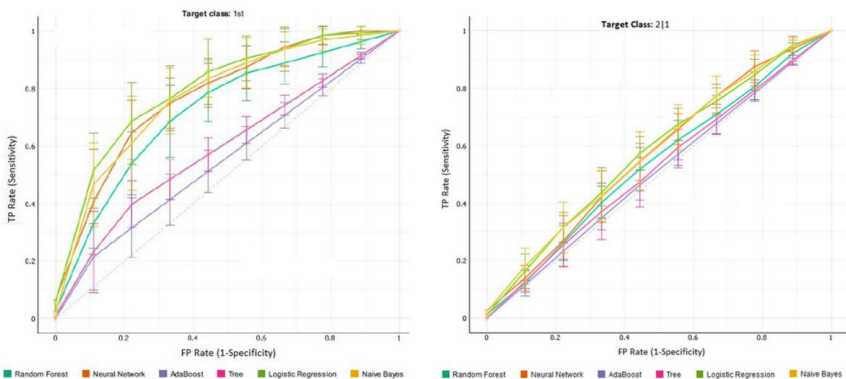


Fig. 9 La sensibilidad del modelo para (a) grado 1st Class (b) grado 2nd

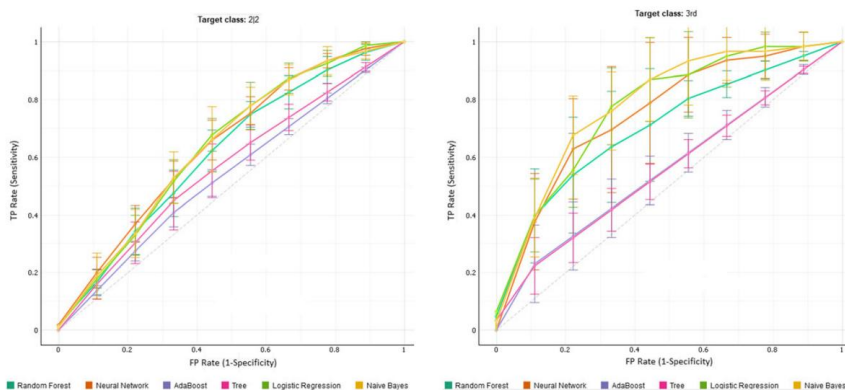


Fig. 10 La sensibilidad del modelo para (a) grado de clase 2|2 (b) grado de 3ra clase

precisión del 51,9 %, seguida de la Regresión Logística con una precisión del 51,1 %, y a esta le sigue el Random Forest con una precisión del 50,3 %. Naïve Bayes obtuvo el cuarto mejor desempeño con una precisión del 50 % y el siguiente en jerarquía es el algoritmo Tree con una precisión del 43,6 %. El algoritmo AdaBoost tuvo la menor precisión con un 42,8 %. La Tabla 4 se obtuvo utilizando el promedio de las clases como la clase objetivo y un tipo de muestreo de validación cruzada estratificado de 10 veces. La Tabla 4 compara los algoritmos de Orange en términos del área bajo la curva ROC (AUC), la precisión de clasificación (CA), la puntuación F1, la tasa de precisión y la recuperación.

La sensibilidad de cada uno de los algoritmos en el modelo de Orange se muestra en las Figs. 9 (a), (b), 10 (a) y 10 (b) utilizando la tasa TP para las cuatro clases de grados, es decir, primera clase, segunda clase inferior, segunda clase superior y tercera clase, respectivamente.

8 Análisis predictivo mediante regresión

Para validar aún más la precisión de los modelos de minería de datos, el conjunto de datos se evaluó más mediante regresión. Se consideraron dos modelos de regresión, los modelos de regresión lineal y de regresión cuadrática, como se presenta en las siguientes secciones. En el modelo de regresión, la relación existente entre los requisitos cognitivos de ingreso (la edad de ingreso del estudiante (x_1), el puntaje agregado WAEC (x_2), el puntaje JAMB (x_3) y el puntaje CUSAS basado en la universidad (x_4) y el CGPA de primer año (y) de los estudiantes fue evaluado.

8.1 Modelo de regresión lineal

El modelo de regresión lineal que representa la relación de los datos analizados se muestra en la ec. 1. Los resultados del modelo se muestran en la Tabla 5, y se puede observar que para los datos de la muestra de 1445 estudiantes, el valor R cuadrado del modelo es 0,207, lo que indica una relación débil. La relación de los parámetros del modelo de regresión lineal se presenta en la ec. 1.

$$y = \frac{1}{4} + x_1 + x_2 + x_3 + x_4$$

819

Tabla 5 Resultados del modelo de regresión lineal

	Estimar	SE	tEstad	pValor
(Interceptar)	1.6187	0.30209	5.3584	9.76E-08
x1	$\bar{y}0.03928$	0.012715	$\bar{y}3.089$	0.002047
x2	0.004029	0.000678	5.9463	3.44E-09
x3	0.010895	0.001817	5.996	2.55E-09
x4	0.3119	0.026534	11.755	1.58E-30

Nº de observaciones: 1445, Grados de libertad de error: 1440

Error RMS: 0,57

R2 : 0,207, R2 ajustado : 0,205

Estadístico F frente a modelo constante: 94,2, valor de p = 3,01E-71

8.2 Modelo de regresión cuadrática

El modelo de regresión cuadrática que representa la relación entre los datos características analizadas se muestra en la ecuación. 2. Los resultados del modelo se muestran en 6 , y se puede observar que para los 1445 datos de la muestra de estudiantes analizados, el valor R-cuadrado es 0,232, lo que también indica una relación débil

Tabla 6 Resultados del modelo de regresión cuadrática

	Estimar	SE	tEstad	pValor
(Interceptar)	4.1111	3.9768	1.0338	0.30141
x1	$\bar{y}0.23418$	0.29097	$\bar{y}0.80484$	0.42105
x2	0.012516	0.014585	0.85812	0.39097
x3	0.007849	0.041166	0.19067	0.84881
x4	$\bar{y}0.61633$	0.54023	$\bar{y}1,1409$	0.25412
x1:x2	0.000208	0.000564	0.36847	0.71258
x1:x3	0.000888	0.00159	0.5588	0.57639
x1:x4	$\bar{y}0.0099$	0.022109	$\bar{y}0.44799$	0.65423
x2:x3	0.000236	7.82E-05	3.0195	0.002577
x2:x4	0.002636	0.00122	2.1617	0.030809
x3:x4	$\bar{y}0.00136$	0.00329	$\bar{y}0,4119$	0.68047
2x1 _	0.003301	0.006568	0.50266	0.61528
2x2 _	$\bar{y}8.20E-05$	2.32E-05	$\bar{y}3.533$	0.000424
2x3 _	$\bar{y}0.00046$	0.000154	$\bar{y}2,9928$	0.002811
2x4 _	0.092437	0.039376	2.3476	0.019033

Nº de observaciones: 1445, Grados de libertad de error: 1430

Error RMS: 0,563

R2 : 0,232, R2 ajustado : 0,224

Estadístico F frente a modelo constante: 30,8, valor p = 4,71E-72

considerando el rango de valores R-cuadrado estándar de 0 a 1. La relación de los parámetros del modelo de regresión cuadrática se presenta en la ec. 2.

$$y = \frac{1}{4} + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} + \beta_{11} x_{11} + \beta_{12} x_{12} + \beta_{13} x_{13} + \beta_{14} x_{14} + \beta_{15} x_{15} + \beta_{16} x_{16} + \beta_{17} x_{17} + \beta_{18} x_{18} + \beta_{19} x_{19} + \beta_{20} x_{20} + \beta_{21} x_{21} + \beta_{22} x_{22} + \beta_{23} x_{23} + \beta_{24} x_{24} + \beta_{25} x_{25} + \beta_{26} x_{26} + \beta_{27} x_{27} + \beta_{28} x_{28} + \beta_{29} x_{29} + \beta_{30} x_{30} + \beta_{31} x_{31} + \beta_{32} x_{32} + \beta_{33} x_{33} + \beta_{34} x_{34} + \beta_{35} x_{35} + \beta_{36} x_{36} + \beta_{37} x_{37} + \beta_{38} x_{38} + \beta_{39} x_{39} + \beta_{40} x_{40} + \beta_{41} x_{41} + \beta_{42} x_{42} + \beta_{43} x_{43} + \beta_{44} x_{44} + \beta_{45} x_{45} + \beta_{46} x_{46} + \beta_{47} x_{47} + \beta_{48} x_{48} + \beta_{49} x_{49} + \beta_{50} x_{50} + \beta_{51} x_{51} + \beta_{52} x_{52} + \beta_{53} x_{53} + \beta_{54} x_{54} + \beta_{55} x_{55} + \beta_{56} x_{56} + \beta_{57} x_{57} + \beta_{58} x_{58} + \beta_{59} x_{59} + \beta_{60} x_{60} + \beta_{61} x_{61} + \beta_{62} x_{62} + \beta_{63} x_{63} + \beta_{64} x_{64} + \beta_{65} x_{65} + \beta_{66} x_{66} + \beta_{67} x_{67} + \beta_{68} x_{68} + \beta_{69} x_{69} + \beta_{70} x_{70} + \beta_{71} x_{71} + \beta_{72} x_{72} + \beta_{73} x_{73} + \beta_{74} x_{74} + \beta_{75} x_{75} + \beta_{76} x_{76} + \beta_{77} x_{77} + \beta_{78} x_{78} + \beta_{79} x_{79} + \beta_{80} x_{80} + \beta_{81} x_{81} + \beta_{82} x_{82} + \beta_{83} x_{83} + \beta_{84} x_{84} + \beta_{85} x_{85} + \beta_{86} x_{86} + \beta_{87} x_{87} + \beta_{88} x_{88} + \beta_{89} x_{89} + \beta_{90} x_{90} + \beta_{91} x_{91} + \beta_{92} x_{92} + \beta_{93} x_{93} + \beta_{94} x_{94} + \beta_{95} x_{95} + \beta_{96} x_{96} + \beta_{97} x_{97} + \beta_{98} x_{98} + \beta_{99} x_{99} + \beta_{100} x_{100}$$

9 Resumen de resultados

En el modelo KNIME, la precisión máxima es del 50,23 % y la precisión más baja es del 39,631 %, mientras que para el modelo Orange la precisión varió del 51,9 % al 42,8 %. El rendimiento de ambos modelos es bastante similar, lo que implica que los algoritmos de minería de datos en ambas plataformas pudieron identificar el alcance de la relación oculta entre los requisitos cognitivos de ingreso a la admisión y la clase de grado de los estudiantes en su primer año. Con una precisión máxima del 51,9 % para ambos modelos de minería de datos; implica que la edad de ingreso de los estudiantes, el puntaje WAEC agregado, el puntaje JAMB y el puntaje CUSAS basado en la universidad, que son los criterios clave de ingreso a los programas 100 L, solo explican parcialmente el desempeño (clase de grado) del estudiante admitido en su primer año en la universidad. Esto significa que otros factores y atributos no académicos de cada estudiante admitido que pueden no ser fáciles de medir en el momento de la admisión, junto con el estilo de vida personal y las luchas mientras están en la universidad son determinantes potenciales del rendimiento académico de los estudiantes en su primer año. año. El modelo de regresión lineal tenía un estadístico F de 94,2 que es significativo con un valor p de 3,01E-71 pero con un valor R2 (coeficiente de determinación) de 0,207; implica que el criterio de entrada de admisión solo explica el 20,7% de la variación del CGPA de 100 L. Un resultado similar se observó para el modelo de regresión cuadrática con un valor de R2 de 0,232, lo que indica una relación predictiva débil.

Los resultados de los modelos de minería de datos KNIME y Orange, y el análisis de regresión realizado en este estudio muestran que existe una relación entre los requisitos de admisión y el rendimiento académico del estudiante admitido en su primer año, pero esta relación no es muy fuerte como lo revelan las precisiones de los algoritmos predictivos. Los criterios de selección para la admisión a la universidad en Nigeria se basan principalmente en el perfil académico de los solicitantes. El resultado de este estudio, por lo tanto, enfatiza la necesidad de reevaluar los criterios de admisión y la necesidad de considerar otros factores no académicos que pueden ser indicadores de la dedicación, motivación y pasión por el éxito de los estudiantes. La excelencia en los deportes, las experiencias de liderazgo, la participación en proyectos estudiantiles a nivel de escuela secundaria, etc. son factores probables que también deben considerarse en el proceso de admisión. Además de los efectos naturales de la destreza académica y las habilidades de aprendizaje de un individuo en su excelencia académica en una universidad, otros factores como el nivel de asistencia a las conferencias, el número de días de enfermedad, los problemas familiares en el hogar que pueden afectar el estado psicológico del estudiante, la economía Los desafíos, los puntajes promedio en el primer conjunto de pruebas, el estilo de vida social, la disponibilidad de programas de apoyo dentro de la universidad, por ejemplo, asesoramiento, etc., también juegan un papel clave en el desempeño general de un estudiante.

En su mayor parte, un proceso de selección bien ejecutado suele producir un buen porcentaje de académicos académicamente sólidos (Idachaba 2018a), pero otros factores como la enseñanza y

la investigación, la calidad de las instalaciones y la gestión institucional general (Idachaba 2018b), pueden afectar en gran medida el rendimiento académico en las universidades. Otros factores que podrían afectar el desempeño académico en las universidades incluyen los cuellos de botella administrativos y la ostentación de las políticas establecidas y las reglas básicas. Garantizar la excelencia académica de los estudiantes en su primer año de estudio requiere un enfoque holístico que abarque la selección de candidatos académicamente prometedores, la provisión de instalaciones, personal y cuerpo docente adecuados; disponibilidad de programas de orientación, asesoramiento y tutoría para estudiantes, disponibilidad de un sistema proactivo de evaluación del desempeño, et

10 Conclusión y alcance futuro

El uso de varios criterios de selección por parte de los organismos reguladores en Nigeria y las universidades nigerianas ha llevado a la admisión de un mayor porcentaje de estudiantes universitarios académicamente sólidos y aspirantes, y a una concentración de grandes talentos en las universidades nigerianas. También ha ayudado a filtrar un mayor porcentaje de candidatos no calificados y reducir la tasa de prácticas de admisión encubiertas en las universidades nigerianas, aunque existen oportunidades existentes para mejorar aún más los procedimientos y prácticas de admisión.

Los requisitos de ingreso para la admisión son índices vitales para seleccionar estudiantes para la admisión en instituciones superiores. El análisis de este estudio revela hasta qué punto los requisitos de ingreso para la admisión se pueden utilizar como predictores del rendimiento una vez que se ingresa al sistema universitario, utilizando la Universidad de Covenant en Nigeria como estudio de caso. Las tendencias de datos presentadas en diagramas de caja muestran que los puntajes de entrada de los estudiantes de WAEC, CUSAS y JAMB son más altos para aquellos que obtuvieron la calificación de 1ra clase, seguidos por aquellos en 2|1, y luego 2|2 y finalmente los estudiantes de 3ra clase tuvieron la menor calificación. puntajes promedio. La extracción de datos educativos realizada con KNIME y la plataforma Orange tuvo una precisión máxima de 50,23 % y 51,9 % respectivamente, y esto muestra que las características cognitivas de ingreso de un estudiante definidas por los requisitos de admisión no explican completamente la eventual clase de primer año. de grado de ese estudiante. Para validar los resultados de los modelos de minería de datos, se desarrollaron modelos analíticos de regresión lineal y cuadrática y se observaron valores de R-cuadrado de 0,207 y 0,232 respectivamente, lo que indica una relación débil entre el CGPA de primer año y los requisitos de ingreso de admisión. Por lo tanto, esto hace que sea vital garantizar que los criterios de selección de admisión sean lo suficientemente sólidos como para permitir una raza bien mezclada de estudiantes con diferentes fortalezas únicas, no necesariamente académicas, que eventualmente puedan traducirse en éxito una vez en la universidad.

En la actualidad, el análisis hasta el momento se ha centrado generalmente en la relación de los criterios de entrada de admisión y el rendimiento de 100 L del estudiante. Será interesante considerar la clasificación de los registros de los estudiantes en función de sus atributos demográficos, como el género, el nivel de ingresos familiares, la religión, el tamaño promedio de la familia, el estado de origen, etc., para identificar nuevas tendencias y conocimientos ocultos. El efecto del enfoque de aula invertida para el aprendizaje de los estudiantes también se puede medir en una sesión académica particular clasificando a los estudiantes después de la admisión para un estudio de casos y controles basado en los criterios de admisión, y luego seleccionando muestras de estudio y de control de cada grupo para someterlas de manera diferente a los modelos de enseñanza tradicionales y de aula invertida. Esto permitirá realizar un análisis comparativo del desempeño del estudio y controlar el CGPA y la clase de grado de los estudiantes al finalizar el ciclo académico.

año, con el fin de identificar si habrá una diferencia significativa entre el desempeño de los estudiantes expuestos al modelo de aula invertida, y aquellos sometidos al modelo de enseñanza tradicional.

Agradecimientos Los autores agradecen al Centro de Investigación, Innovación y Desarrollo de la Universidad Covenant por su compromiso con la investigación innovadora y por proporcionar un entorno propicio para la investigación.

Nota del editor Springer Nature se mantiene neutral con respecto a los reclamos jurisdiccionales en mapas publicados y afiliaciones institucionales.

Referencias

- Adeogun, A., Subair, S. y Osifila, G. (2009). Desregulación de la educación universitaria en Nigeria: problemas y perspectivas. *Florida Journal of Educational Administration & Policy*, 3, 1–8.
- Adeyemi, K. (2001). La igualdad de acceso y el factor del área de captación en las admisiones universitarias en Nigeria. *Más alto Educación*, 42, 307–332.
- Agarwal, S., Pandey, G. y Tiwari, M. (2012). Minería de datos en educación: clasificación de datos y enfoque de árboles de decisión. *Revista internacional de educación electrónica, comercio electrónico, gestión electrónica y aprendizaje electrónico*, 2, 140.
- Ahmed, ABED y Elaraby, IS (2014). Minería de datos: una predicción para el rendimiento del estudiante utilizando el método de clasificación. *Revista mundial de aplicaciones informáticas y tecnología*, 2, 43–47.
- Ahuja, R. y Kankane, Y. (2018). Predecir la probabilidad de que un estudiante complete su título usando diferentes técnicas de minería de datos. 474–477.
- Ahuja, R., Jha, A., Maurya, R. y Srivastava, R. (2019). Análisis de minería de datos educativos. 4.ª Conferencia internacional sobre búsqueda de armonía, computación blanda y aplicaciones, ICHSA 2018, Gurgaon; India. *Avances en sistemas inteligentes y computación*, 897–907.
- Aina, OI 2002. Modos alternativos de financiación de la educación superior en Nigeria y las implicaciones para gobierno universitario. *África Development/Afrique et Développement*, 236–262.
- Ajadi, TO, Salawu, IO y Adeoye, FA (2008). E-learning y educación a distancia en Nigeria. En línea Sumisión, 7.
- Ajayi, I. y Ekundayo, HT (2008). La desregulación de la educación universitaria en Nigeria: Implicaciones para seguro de calidad. *Nebulosa*, 5, 212–224.
- Almarabeh, H. (2017). Análisis del desempeño de los Estudiantes mediante el uso de diferentes clasificadores de minería de datos. *Revista Internacional de Educación Moderna y Ciencias de la Computación*, 9, 9.
- Aluede, O., Idogho, PO & Imonikhe, JS (2012). Aumentar el acceso a la educación universitaria en Nigeria: Presentar retos y sugerencias para el futuro. *El Simposio Africano*. 3–13.
- Angeli, C., Howard, SK, Ma, J., Yang, J. y Kirschner, PA (2017). Minería de datos en la investigación en el aula de tecnología educativa: ¿Puede hacer una contribución? *Informática y Educación*, 113, 226–242.
- Arsad, PM y Buniyamin, N. (2014). Métodos de Redes Neuronales y Regresión Lineal para la predicción del rendimiento académico de los estudiantes. *Conferencia Global de Educación en Ingeniería (EDUCON)*, IEEE, 2014. IEEE, 916–921.
- Atta UR, R., Sultan, K., Aldhafferi, N. y Alqahtani, A. (2018). Minería de datos educativos para mejorar la enseñanza y el aprendizaje. *Revista de tecnología de la información teórica y aplicada*, 96, 4417–4427.
- Azevedo, A. (2018). Minería de datos y descubrimiento de conocimiento en bases de datos. En *Enciclopedia de la ciencia y la tecnología de la información*, cuarta edición. IGI Global.
- Babalola, JB (1998). Costo y financiamiento de la educación universitaria en Nigeria. *Educación Superior*, 36, 43–66.
- Baepler, P. y Murdoch, CJ (2010). Analítica académica y minería de datos en la educación superior. *Internacional Revista para la Beca de Enseñanza y Aprendizaje*, 4, 17.
- Bharara, S., Sabitha, S. y Bansal, A. (2018). Aplicación de análisis de aprendizaje utilizando minería de datos de agrupamiento para el análisis de disposición de los estudiantes. *Educación y tecnologías de la información*, 23, 957–984.
- Bhise, R., Thorat, S. y Supekar, A. (2013). Importancia de la minería de datos en el sistema de educación superior. *IOSR Revista de Humanidades y Ciencias Sociales (IOSR-JHSS)*, 6, 18.
- Bucos, M. y Drăgulescu, B. (2018). Predecir el éxito de los estudiantes utilizando datos generados en la educación tradicional entornos. *Revista TEM*, 7, 617–625.

- Burgos, C., Campanario, ML, Peña, DDL, Lara, JA, Lizcano, D., & Martínez, MA (2018). Minería de datos para modelar el desempeño de los estudiantes: un plan de acción de tutoría para prevenir la deserción académica. *Informática e ingeniería eléctrica*, 66, 541–556.
- Burke, AS y Fedorek, B. (2017). ¿Bflipping[^] promueve el compromiso?: Una comparación de un tradicional, en línea y clase invertida. *Aprendizaje activo en la educación superior*, 18, 11–24.
- Daradouis, T., Marqués Puig, JM, Arguedas, M., & Calvet Liñan, L. (2019). Analizar las percepciones de los estudiantes para mejorar el diseño de una herramienta de evaluación automatizada en programación distribuida en línea. *Informática y Educación*, 128, 159–170.
- Ebbeler, J. (2013). Introducción a la antigua Roma', la versión invertida. *The Chronicle of Higher Education* [en línea], 59. Disponible: <https://www.chronicle.com/article/Introduction-to-Ancient/140475/>.
- Fernandes, E., Holanda, M., Victorino, M., Borges, V., Carvalho, R., & Erven, GV (2019). Minería de datos educativos: análisis predictivo del rendimiento académico de estudiantes de escuelas públicas de la capital de Brasil. *Revista de Investigación Empresarial*, 94, 335–343.
- Hussain, S., Dahan, NA, Ba-Alwib, FM y Ribata, N. (2018). Minería de datos educativos y análisis del rendimiento académico de los estudiantes utilizando WEKA. *Revista indonesia de ingeniería eléctrica y ciencias de la computación*, 9, 447–459.
- Hussain, S., Atallah, R., Kamsin, A. y Hazarika, J. (2019). Regla de clasificación, agrupamiento y asociación minería en conjuntos de datos educativos utilizando herramientas de minería de datos: un estudio de caso.
- Ibrahim, ZM, Bader-El-Den, M. y Cocea, M. (2019). Retroalimentación de la unidad de minería para explorar el aprendizaje de los estudiantes experiencias. *Avances en Sistemas Inteligentes y Computación*.
- Idachaba, F. (2018a). Desarrollo de un esquema de tutoría rápida para la gestión de clases grandes en departamentos de ingeniería. *Conferencia INTED2018*, 5 al 7 de marzo de 2018 2018a Valencia, España. 5725–5731.
- Idachaba, F. (2018b). Diseño de currículo de ingeniería basado en resultados: un sistema para simplificar el currículo y mejorar la calidad de los graduados en ingeniería. *Conferencia INTED2018*, 5 al 7 de marzo de 2018 2018b Valencia, España. 5888–5893.
- Kaur, N. y Kaur, J. (2018). Evaluación del desempeño de la clasificación de minería de datos en el sistema educativo usando algoritmo genético. *Revista internacional de ciencia y tecnología avanzadas*, 114, 127–138.
- Khan, A. y Ghosh, SK (2018). Análisis basado en minería de datos para explorar el efecto de la enseñanza en los estudiantes rendimiento. *Educación y tecnologías de la información*, 23, 1677–1697.
- Khedr, AE y El Seddawy, AI (2015). Un marco de minería de datos propuesto para el sistema de educación superior. *Revista internacional de aplicaciones informáticas*, 113, 24–31.
- Kim, D., Yoon, M., Jo, IH y Branch, RM (2018). Análisis de aprendizaje para apoyar el aprendizaje autorregulado en cursos en línea asincrónicos: un estudio de caso en una universidad de mujeres en Corea del Sur. *Informática y Educación*, 127, 233–251.
- Kostopoulos, G., Kotsiantis, S., Pierrakeas, C., Koutsonikos, G. y Gravvanis, GA (2018). Pronosticar el éxito de los estudiantes en una universidad abierta. *Revista internacional de tecnología de aprendizaje*, 13, 26–43.
- Kurt, G. (2017). Implementando el aula invertida en la formación docente: Evidencia de Turquía. *Educativo Tecnología y Sociedad*, 20, 211–221.
- Lynch, CF (2017). ¿Quiénes profetizan desde el big data en la educación? Nuevos conocimientos y nuevos desafíos. *Teoría y Investigación en Educación*, 15, 249–271.
- Nikolic, S. y Nicholls, B. (2018). Explorar el interés de los estudiantes por el aprendizaje asistido por pares en línea utilizando tecnología de la realidad.
- Nurhayati, OD, Bachri, OS, Supriyanto, A. y Hasbullah, M. (2018). Sistema de predicción de graduación mediante red neuronal artificial. *Revista internacional de ingeniería mecánica y tecnología*, 9, 1051– 1057.
- Odukoya, JA, Popoola, SI, Atayero, AA, Omole, DO, Badejo, JA, John, TM y Olowo, OO (2018). Análisis de aprendizaje: conjunto de datos para la evaluación empírica de los requisitos de ingreso a los programas de pregrado de ingeniería en una universidad nigeriana. *Resumen de datos*, 17, 998–1014.
- Oguntunde, P., Okagbue, H., Oguntunde, OA y Opanuga, A. (2018). Análisis de la interrelación entre los resultados del primer año de los estudiantes y sus notas finales de graduación. *Revista Internacional de Ciencias Avanzadas y Aplicadas*, 5, 1–6.
- Olsson, M. y Mozelius, P. (2016). Sobre el diseño de entornos de aprendizaje en línea para la enseñanza de la programación. *Actas de la Conferencia Europea sobre e-Learning, ECEL*. 533–539.
- Osmanbegović, E. y Suljić, M. (2012). Enfoque de minería de datos para predecir el rendimiento de los estudiantes.
- Popoola, SI, Atayero, AA, Badejo, JA, Odukoya, JA, Omole, DO y Ajayi, P. (2018). Conjuntos de datos sobre tendencias demográficas en la inscripción en programas de pregrado de ingeniería en la Universidad de Covenant, Nigeria. *Resumen de datos*, 18, 47–59.
- Rodrigues, MW, Isotani, S., & Zárata, LE (2018). Minería de datos educativos: una revisión del proceso de evaluación en el e-learning. *Telemática e informática*, 35, 1701–1717.

- Roy, S. y Garg, A. (2018). Predecir el rendimiento académico de los estudiantes utilizando técnicas de clasificación. 2017 4ª conferencia internacional de la sección IEEE Uttar Pradesh sobre electricidad, informática y electrónica, UPCON 2017. 568–572.
- Ryan, SJD y Baker. (2010). Minería de datos para la educación. *Enciclopedia internacional de educación*, 7, 112–118.
- Saint, W., Hartnett, TA y Strassner, E. (2003). Educación superior en Nigeria: un informe de estado. *Política de educación superior*, 16, 259–281.
- Senthil, S. y Lin, WM (2018). Aplicar técnicas de clasificación para predecir los resultados académicos de los alumnos. Conferencia internacional IEEE de 2017 sobre tendencias actuales en computación avanzada, ICCTAC 2017. 1–6.
- Sivakumar, S. & Selvaraj, R. 2018. Modelado predictivo del desempeño de los estudiantes a través de la mejora árbol de decisión. *Apuntes de clase en ingeniería eléctrica*.
- Tair, MMA y El-Halees, AM (2012). Minería de datos educativos para mejorar el rendimiento de los estudiantes: un estudio de caso. *Revista internacional de investigación de tecnologías de la información y la comunicación*, 2, 140–146.
- Van WYK, MM (2018). Opiniones de estudiantes de economía sobre la utilidad de un enfoque pedagógico de aula invertida para un entorno de aprendizaje electrónico abierto a distancia. *Revista internacional de tecnología de la información y el aprendizaje*, 35, 255–265.
- Yadav, SK, Bharadwaj, B. y PAL, S. (2012). Minería de datos educativos para predecir la retención de los estudiantes: A estudio comparativo. preimpresión de arXiv arXiv:1203.2987.