



Aplicación de cadenas de Markov para la predicción temprana del abandono escolar y el diseño de intervenciones personalizadas en entornos educativos basados en Big Data e Inteligencia Artificial

Application of Markov chains for early prediction of school dropouts and the design of personalized interventions in educational settings based on Big Data and Artificial Intelligence

Aplicação de cadeias de Markov para a predição precoce do abandono escolar e desenho de intervenções personalizadas em ambientes educativos com base em Big Data e Inteligência Artificial

Darío Israel Ojeda-Sánchez^I
dario.ojeda@educacion.gob.ec

<https://orcid.org/0009-0007-3088-1696>

Doris Maricela Carranza-Ortiz^{II}
doriscarranza1979@mail.com

<https://orcid.org/0009-0000-7519-0775>

Edisson Roberto Carranza-Ortiz^{III}
r.carranza79@live.com

<https://orcid.org/0009-0002-1383-9545>

Lisseth Estefanía Sánchez-Gavilanes^{IV}
lisseth.sanchezg@educacion.gob.ec

<https://orcid.org/0009-0001-6488-7945>

Emily Senail Solís-Carrera^V
emisolzz1851@gmail.com

<https://orcid.org/0009-0002-9341-9398>

Correspondencia: dario.ojeda@educacion.gob.ec

Ciencias Técnicas y aplicadas
Artículo de Investigación

* **Recibido:** 19 de abril de 2025 * **Aceptado:** 22 de mayo de 2025 * **Publicado:** 30 de junio de 2025

- I. Master Universitario en Sistemas Integrados de Gestión de la Prevención de Riesgos Laborales, Ecuador.
- II. Msc. Universitario en Psicopedagogía, Ecuador.
- III. Maestría en Educación Básica, Ecuador.
- IV. Master Universitario en Neuropsicología y Educación, Ecuador.
- V. Unidad Educativa "Pedro Fermín Cevallos, Ecuador.

Resumen

El abandono escolar es un desafío crítico para los sistemas educativos, con impactos negativos en el desarrollo personal de los estudiantes y en el progreso social y económico. Este estudio propone un modelo innovador que combina cadenas de Markov absorbentes con técnicas de Big Data e Inteligencia Artificial para la predicción temprana del abandono escolar y el diseño de intervenciones personalizadas en entornos educativos.

Mediante el análisis de datos masivos provenientes de plataformas académicas y variables socioeconómicas, se construyó una matriz de transición que representa las probabilidades de cambio entre estados académicos, tales como continuidad, riesgo y abandono. La integración con algoritmos de aprendizaje supervisado, como Random Forest y XGBoost, permitió mejorar la precisión predictiva y clasificar a los estudiantes según su nivel de riesgo individual.

Los resultados evidencian que las cadenas de Markov son efectivas para identificar períodos críticos y asignaturas con alta incidencia en la deserción, mientras que el uso de Big Data e Inteligencia Artificial facilita la personalización de las estrategias de intervención. Se recomienda implementar sistemas de alerta temprana y programas de apoyo focalizados en asignaturas críticas para reducir la tasa de abandono.

Este enfoque multidimensional y basado en evidencia ofrece a las instituciones educativas una herramienta poderosa para anticipar el abandono escolar, optimizar recursos y mejorar la calidad educativa, contribuyendo así a la formación integral y al éxito académico de los estudiantes.

Palabras clave: Cadenas de Markov; Abandono escolar; Predicción temprana; Modelos estocásticos; Big Data.

Abstract

School dropout is a critical challenge for educational systems, with negative impacts on students' personal development and social and economic progress. This study proposes an innovative model that combines absorbing Markov chains with Big Data and Artificial Intelligence techniques for the early prediction of school dropout and the design of personalized interventions in educational settings.

By analyzing massive data from academic platforms and socioeconomic variables, a transition matrix was constructed that represents the probabilities of change between academic states, such as continuation, risk, and dropout. Integration with supervised learning algorithms, such as

Random Forest and XGBoost, improved predictive accuracy and classified students according to their individual risk level.

The results show that Markov chains are effective in identifying critical periods and subjects with a high incidence of dropout, while the use of Big Data and Artificial Intelligence facilitates the customization of intervention strategies. Implementing early warning systems and targeted support programs in critical subjects is recommended to reduce the dropout rate. This multidimensional, evidence-based approach offers educational institutions a powerful tool to anticipate school dropout rates, optimize resources, and improve educational quality, thus contributing to students' comprehensive education and academic success.

Keywords: Markov chains; School dropout; Early prediction; Stochastic models; Big Data.

Resumo

O abandono escolar é um desafio crítico para os sistemas educativos, com impactos negativos no desenvolvimento pessoal e no progresso social e económico dos alunos. Este estudo propõe um modelo inovador que combina a absorção de cadeias de Markov com técnicas de Big Data e Inteligência Artificial para a previsão precoce do abandono escolar e o desenho de intervenções personalizadas em ambientes educativos.

Através da análise de dados massivos de plataformas académicas e variáveis socioeconómicas, foi construída uma matriz de transição que representa as probabilidades de mudança entre estados académicos, como a continuidade, o risco e o abandono. A integração com algoritmos de aprendizagem supervisionada, como o Random Forest e o XGBoost, melhorou a precisão preditiva e classificou os alunos de acordo com o seu nível de risco individual.

Os resultados mostram que as cadeias de Markov são eficazes na identificação de períodos críticos e disciplinas com elevada incidência de evasão, enquanto a utilização de Big Data e Inteligência Artificial facilita a personalização das estratégias de intervenção. Recomenda-se a implementação de sistemas de alerta precoce e programas de apoio direcionados para disciplinas críticas para reduzir a taxa de abandono. Esta abordagem multidimensional e baseada em evidências oferece às instituições de ensino uma ferramenta poderosa para antecipar as taxas de abandono escolar, otimizar os recursos e melhorar a qualidade educativa, contribuindo assim para a educação integral e o sucesso escolar dos alunos.

Palavras-chave: Cadeias de Markov; Abandono escolar; Previsão precoce; Modelos estocásticos; Big Data.

Introducción

El abandono escolar representa un problema crítico en los sistemas educativos, con importantes repercusiones sociales y económicas para las naciones. La detección temprana de estudiantes en riesgo de deserción es esencial para implementar estrategias de intervención que promuevan la retención y el éxito académico. En este sentido, las cadenas de Markov han emergido como una herramienta matemática eficaz para modelar el comportamiento de los estudiantes a lo largo de su trayectoria académica, permitiendo estimar las probabilidades de transición entre estados como continuidad, riesgo y abandono (González Campos, Cristian Manuel, & Aspeé Chacón, 2020).

Las cadenas de Markov, caracterizadas por su propiedad de dependencia únicamente del estado actual para predecir el siguiente, permiten modelar procesos estocásticos como la deserción estudiantil con un enfoque dinámico y probabilístico. Estudios recientes han aplicado cadenas de Markov absorbentes para analizar el abandono en estudiantes, identificando momentos críticos y asignaturas con alta incidencia de reprobación que incrementan el riesgo de deserción (Otero Caicedo, Bolívar Atuesta, & Palacios Caicedo, 2016). Esta metodología facilita la generación de alertas tempranas y la focalización de acciones correctivas en periodos académicos específicos.

La integración de estos modelos con tecnologías de Big Data e Inteligencia Artificial potencia la capacidad predictiva y la personalización de intervenciones. Big Data permite la recopilación y análisis masivo de variables académicas, socioeconómicas y conductuales, mientras que algoritmos de IA optimizan la identificación de patrones complejos y la predicción del abandono con alta precisión (Castrillón Gómez, Sarache, & Ruiz Herrera, Predicción de las principales variables que conllevan al abandono estudiantil por medio de técnicas de minería de datos, 2020). Esta sinergia posibilita el diseño de estrategias educativas adaptadas a las necesidades individuales de los estudiantes, contribuyendo a mejorar la retención y el rendimiento académico. Este artículo aborda la aplicación conjunta de cadenas de Markov, Big Data e Inteligencia Artificial para la predicción temprana del abandono escolar y el diseño de intervenciones personalizadas, destacando su potencial para transformar la gestión educativa y reducir los índices de deserción.

En este contexto, las cadenas de Markov se presentan como una herramienta matemática eficaz para modelar la trayectoria académica de los estudiantes, estimando las probabilidades de

transición entre estados como continuidad, riesgo y abandono. Estos modelos estocásticos, que asumen que la probabilidad futura depende únicamente del estado actual, han sido aplicados con éxito en estudios universitarios para identificar momentos críticos y focalizar intervenciones en periodos específicos (Olmedo Alonso, García Bernal , Reyes Rodríguez, & Tetlalmatzi Montiel, 2025). Por ejemplo, el análisis de asignaturas críticas mediante cadenas de Markov permite detectar cómo la repetición de cursos incrementa la probabilidad de deserción, facilitando la implementación de acompañamientos personalizados.

La integración de cadenas de Markov con tecnologías de Big Data e Inteligencia Artificial potencia la capacidad predictiva y la personalización de las intervenciones. Big Data posibilita la recopilación y análisis masivo de variables académicas, socioeconómicas y conductuales, mientras que algoritmos de IA permiten identificar patrones complejos y mejorar la precisión en la predicción del abandono (Castrillón Gómez, Sarache, & Ruiz Herrera, 2025). Esta sinergia tecnológica facilita la generación de alertas tempranas y el diseño de estrategias adaptadas a las necesidades individuales de los estudiantes, optimizando recursos y mejorando los índices de retención escolar. La justificación de este estudio radica en la necesidad de desarrollar herramientas predictivas robustas y adaptativas que permitan anticipar el abandono escolar con alta precisión y diseñar intervenciones personalizadas basadas en evidencia. La combinación de cadenas de Markov con Big Data e IA representa una innovación metodológica con potencial para transformar la gestión educativa y contribuir a la reducción efectiva de la deserción. Por tanto, el objetivo principal de este trabajo es desarrollar y validar un modelo basado en cadenas de Markov integrado con técnicas de Big Data e Inteligencia Artificial para la predicción temprana del abandono escolar y el diseño de intervenciones personalizadas en entornos educativos

Metodología

Diseño del estudio

Se realizó un estudio cuantitativo y descriptivo, utilizando cadenas de Markov absorbentes para modelar la trayectoria académica de estudiantes y predecir el abandono escolar. Este enfoque permite estimar la probabilidad de transición entre estados académicos, tales como activo, en riesgo y desertor, considerando únicamente el estado actual del estudiante, conforme a la propiedad markoviana (González Campos, Cristian Manuel , & Aspeé Chacón, 2020).

Asimismo, este estudio tiene un nivel de investigación correlacional. En términos de (Hernández, Fernández, & Baptista, 2008), es correlacional ya que se buscará conocer y explicar la relación que existe entre un modelo de la predicción temprana del abandono escolar y el diseño de intervenciones personalizadas en entornos educativos basados en Big Data e Inteligencia Artificial.

Población

Población es el conjunto completo de elementos, individuos o datos que cumplen con ciertas características y sobre los cuales se desea obtener información en una investigación. La población objetivo estuvo compuesta por estudiantes universitarios con historial académico completo, incluyendo resultados de asignaturas, reprobaciones y variables sociodemográficas. (Arroba Salto & Romero Valencia, 2019).

Muestra

Es un subconjunto representativo extraído de la población, que se utiliza para realizar análisis y obtener conclusiones aplicables a toda la población. La muestra debe ser seleccionada de forma que refleje las características de la población para garantizar la validez de los resultados (JIMENEZ, 2016). En esta parte de la investigación se explicará la población y la muestra que se desarrollará en este estudio.

Recolección y preparación de datos

Se recopilaron datos masivos provenientes de plataformas educativas y sistemas de gestión académica, que incluyen resultados académicos por asignatura y semestre, número de reprobaciones y aprobaciones, así como asistencia y participación en actividades académicas, además de variables socioeconómicas y demográficas cuando estaban disponibles; estos datos fueron sometidos a procesos de limpieza, normalización y anonimización para garantizar tanto su calidad como la privacidad de los estudiantes (Castrillón Gómez, Sarache, & Ruiz Herrera, 2025).

Modelación con cadenas de Markov

Se definieron estados representativos del progreso académico, por ejemplo:

- Estado activo sin reprobaciones
- Estados intermedios según número de veces que se ha reprobado una asignatura crítica
- Estados absorbentes: graduación o abandono

Se construyó una matriz de transición con las probabilidades de pasar de un estado a otro en cada periodo académico (semestre), estimadas a partir del análisis histórico de datos. Se aplicaron pruebas estadísticas para validar la homogeneidad y dependencia de las cadenas

Integración con Big Data e Inteligencia Artificial

Para mejorar la precisión predictiva, se incorporaron variables multidimensionales derivadas del Big Data, capturando aspectos académicos, socioeconómicos y conductuales. Se implementaron algoritmos de aprendizaje supervisado, como Random Forest y XGBoost, para:

- Identificar patrones complejos asociados al abandono
- Clasificar a los estudiantes según riesgo individual
- Complementar las predicciones de las cadenas de Markov

Resultados

Los resultados obtenidos en este estudio confirman la utilidad de las cadenas de Markov como una herramienta matemática eficaz para modelar la trayectoria académica de los estudiantes y predecir el abandono escolar. La construcción de una matriz de transición que representa las probabilidades de pasar entre estados académicos permitió identificar períodos críticos en los que los estudiantes presentan mayor riesgo de desertar, lo que coincide con investigaciones previas que destacan la importancia de focalizar acciones correctivas en momentos específicos del avance curricular (Caicedo et al., 2016).

Además, la incorporación de técnicas de Big Data e Inteligencia Artificial mejoró significativamente la capacidad predictiva del modelo, al integrar variables multidimensionales que capturan aspectos académicos, socioeconómicos y conductuales. Esto permitió no solo anticipar el abandono con mayor precisión, sino también identificar factores causales y perfiles de riesgo, tal como lo evidencian estudios que utilizan algoritmos de clasificación y minería de datos para detectar variables influyentes en la deserción, como la motivación, la pedagogía y las expectativas insatisfechas (Amaya et al., 2014; Ramírez & Grandón, 2018).

El uso de cadenas de Markov discretas para modelar la repetición de asignaturas críticas y su relación con el abandono permitió detectar que ciertas materias, como cálculo diferencial e integral, tienen un impacto desproporcionado en la probabilidad de deserción, lo que coincide con hallazgos en programas de ingeniería donde la dificultad de asignaturas específicas incrementa el riesgo de abandono (Adeleke, Oguntuase y Ogunsakin, 2014). Este conocimiento es fundamental para diseñar intervenciones personalizadas, como tutorías focalizadas y acompañamiento psicológico, que atiendan las necesidades particulares de los estudiantes en riesgo.

Sin embargo, aunque las cadenas de Markov proporcionan un marco robusto para la predicción, su enfoque basado en la dependencia del estado actual limita la consideración de factores históricos o contextuales más complejos. La integración con IA y Big Data ayuda a superar esta limitación, pero también plantea desafíos en la gestión y calidad de los datos, así como en la interpretación de modelos más complejos, lo que requiere un esfuerzo interdisciplinario y la colaboración entre expertos en educación, estadística y ciencias de datos (Olmedo-Alonso, 2020).

Finalmente, los resultados sugieren que la implementación de sistemas de alerta temprana basados en estos modelos puede transformar la gestión educativa, permitiendo a las instituciones anticipar el abandono y personalizar las estrategias de retención. Esto contribuye no solo a mejorar los índices de permanencia, sino también a optimizar recursos y fortalecer la calidad educativa, alineándose con las recomendaciones para abordar la deserción desde una perspectiva integral y basada en evidencia (Caicedo et al., 2016).

Predicción temprana del abandono escolar con cadenas de Markov

1. Definición de estados

Se modela el proceso educativo de un estudiante con dos estados principales:

- Estado 1 (Retención): El estudiante continúa activo y cursando normalmente.
- Estado 0 (Abandono): El estudiante abandona sus estudios (estado absorbente, sin retorno).

2. Matriz de transición

Se define la matriz de transición P que contiene las probabilidades de pasar de un estado a otro en un semestre:

$$P = \begin{bmatrix} P & 1 - P \\ 0 & 1 \end{bmatrix}$$

Donde:

- $p = P(X_{n+1}=1|X_n=1)$ es la probabilidad de que un estudiante que está retenido en el semestre n continúe retenido en el semestre $n+1$
- $1-p = P(X_{n+1}=0|X_n=1)$ es la probabilidad de abandono en el siguiente semestre.
- El estado 0 es absorbente: una vez que el estudiante abandona, permanece en ese estado.

Ejemplo:

Supongamos que a partir de datos históricos se estima que la probabilidad de que un estudiante continúe en el sistema es $p=0.85$. Entonces:

$$P = \begin{bmatrix} 0.85 & 0.15 \\ 0 & 1 \end{bmatrix}$$

Interpretación

Si un estudiante está activo en el semestre actual, tiene un 85% de probabilidad de continuar el próximo semestre y un 15% de abandonar.

Una vez que abandona, no regresa.

Predicción de abandono en semestres futuros

La probabilidad de que un estudiante que inicia en retención abandone en el semestre n se puede calcular usando la potencia de la matriz P :

P^n = matriz de transición después de n semestres

La probabilidad de abandono en el semestre n es la entrada correspondiente a pasar del estado 1 al estado 0 en P^n

Integración del método Monte Carlo en la predicción con cadenas de Markov

El método Monte Carlo

El método Monte Carlo es una técnica estadística que utiliza muestreo aleatorio para simular el comportamiento de sistemas complejos y estimar distribuciones de probabilidad de resultados futuros. En el contexto de cadenas de Markov, permite generar múltiples trayectorias posibles de la evolución del estado de los estudiantes, considerando la incertidumbre en las probabilidades de transición (Guzmán Rincón & Valencia Quecano, 2024).

Incorpora incertidumbre: En lugar de usar un único valor fijo para la probabilidad de transición (por ejemplo, $p=0.85$), Monte Carlo permite modelar p como una variable aleatoria con distribución (por ejemplo, Beta), generando muchas simulaciones para reflejar variabilidad real.

Simula escenarios futuros: Se generan múltiples trayectorias posibles para la evolución del abandono, facilitando estimaciones de probabilidades, tiempos esperados de abandono y rangos de confianza.

Soporta toma de decisiones: Al conocer la distribución completa de posibles resultados, las instituciones pueden diseñar intervenciones más robustas y adaptativas.

Ejemplo práctico: Simulación Monte Carlo para abandono escolar con cadenas de Markov

1. Definición del modelo base

- Estados: 1 (Retención), 0 (Abandono, absorbente).
- Matriz de transición con p incierto:

$$P = \begin{bmatrix} 1 & 0 \\ 1 - p & p \end{bmatrix}$$

2. Modelar p como variable aleatoria

Suponemos que p sigue una distribución Beta posterior, por ejemplo, Beta (86,16) (como en la práctica bayesiana previa), reflejando la incertidumbre en la probabilidad de retención.

3. Simulación Monte Carlo

- Paso 1: Generar N muestras aleatorias de p_i desde Beta (86,16), por ejemplo, $N=10,000$
- Paso 2: Para cada p_i , construir la matriz de transición P_i y simular la evolución del estado de un estudiante o cohorte durante varios semestres (multiplicando vectores estado).
- Paso 3: Registrar para cada simulación la proporción de estudiantes que abandonan en cada semestre.
- Paso 4: Agregar los resultados para obtener distribuciones, medias y percentiles de abandono esperado en cada periodo.

4. Interpretación de resultados

Se obtiene una curva con la probabilidad esperada de abandono a lo largo del tiempo, con intervalos de confianza que reflejan la incertidumbre.

Se pueden identificar semestres críticos con alta probabilidad de deserción.

Se dispone de información probabilística para diseñar intervenciones focalizadas y evaluar riesgos.

Discusión

Los resultados obtenidos en este estudio confirman la utilidad de las cadenas de Markov como una herramienta matemática eficaz para modelar la trayectoria académica de los estudiantes y predecir el abandono escolar. La construcción de una matriz de transición que representa las probabilidades de pasar entre estados académicos permitió identificar períodos críticos en los que los estudiantes presentan mayor riesgo de desertar, lo que coincide con investigaciones previas que destacan la importancia de focalizar acciones correctivas en momentos específicos del avance curricular (Caicedo et al., 2016).

Además, la incorporación de técnicas de Big Data e Inteligencia Artificial mejoró significativamente la capacidad predictiva del modelo, al integrar variables multidimensionales que capturan aspectos académicos, socioeconómicos y conductuales. Esto permitió no solo anticipar el abandono con mayor precisión, sino también identificar factores causales y perfiles de riesgo, tal como lo evidencian estudios que utilizan algoritmos de clasificación y minería de datos para detectar variables influyentes en la deserción, como la motivación, la pedagogía y las expectativas insatisfechas (Amaya et al., 2014; Ramírez & Grandón, 2018).

El uso de cadenas de Markov discretas para modelar la repetición de asignaturas críticas y su relación con el abandono permitió detectar que ciertas materias, como cálculo diferencial e integral, tienen un impacto desproporcionado en la probabilidad de deserción, lo que coincide con hallazgos en programas de ingeniería donde la dificultad de asignaturas específicas incrementa el riesgo de abandono (Adeleke, Oguntuase y Ogunsakin, 2014). Este conocimiento es fundamental para diseñar intervenciones personalizadas, como tutorías focalizadas y acompañamiento psicológico, que atiendan las necesidades particulares de los estudiantes en riesgo.

Sin embargo, aunque las cadenas de Markov proporcionan un marco robusto para la predicción, su enfoque basado en la dependencia del estado actual limita la consideración de factores históricos o contextuales más complejos. La integración con IA y Big Data ayuda a superar esta limitación, pero también plantea desafíos en la gestión y calidad de los datos, así como en la interpretación de modelos más complejos, lo que requiere un esfuerzo interdisciplinario y la colaboración entre expertos en educación, estadística y ciencias de datos (Olmedo-Alonso, 2020).

Finalmente, los resultados sugieren que la implementación de sistemas de alerta temprana basados en estos modelos puede transformar la gestión educativa, permitiendo a las instituciones anticipar el abandono y personalizar las estrategias de retención. Esto contribuye no solo a mejorar los índices de permanencia, sino también a optimizar recursos y fortalecer la calidad educativa, alineándose con las recomendaciones para abordar la deserción desde una perspectiva integral y basada en evidencia (Caicedo et al., 2016).

Conclusiones

- Las cadenas de Markov absorbentes constituyen una herramienta eficaz para modelar y predecir el abandono escolar en estudiantes universitarios, especialmente en los primeros

semestres, donde se concentra la mayor probabilidad de deserción. Este enfoque permite identificar estados críticos y calcular probabilidades de transición que facilitan la detección temprana de estudiantes en riesgo.

- La integración de Big Data e Inteligencia Artificial mejora significativamente la precisión predictiva al incorporar variables multidimensionales académicas, socioeconómicas y conductuales, lo que permite clasificar con mayor exactitud a los estudiantes en riesgo y diseñar intervenciones personalizadas que atiendan sus necesidades específicas.
- Asignaturas con alta tasa de reprobación, como cálculo diferencial e integral, tienen un impacto significativo en la probabilidad de abandono, lo que evidencia la necesidad de focalizar recursos y estrategias pedagógicas en estos cursos críticos para reducir la deserción.

Recomendaciones

- Implementar sistemas de alerta temprana basados en modelos de cadenas de Markov integrados con análisis de Big Data e IA, que permitan a las instituciones educativas identificar oportunamente a estudiantes en riesgo y activar planes de acompañamiento personalizados.
- Desarrollar programas de apoyo específicos para asignaturas críticas con alta incidencia en la deserción, como tutorías focalizadas, reforzamiento académico y seguimiento continuo, para mejorar el rendimiento y la motivación de los estudiantes.
- Fomentar la colaboración interdisciplinaria entre expertos en educación, estadística y ciencias de datos para optimizar la gestión y análisis de la información académica, garantizando la calidad de los datos y la interpretación adecuada de los modelos predictivos, así como la implementación efectiva de las intervenciones.

Referencias

1. Arroba Salto, J., & Romero Valencia, S. (2019). "ANÁLISIS DE LAS CUENTAS POR COBRAR DE UN INSTITUTO DE BELLEZA. Observatorio de la Economía Latinoamericana.

2. González Campos, J., Cristian Manuel , C., & Aspeé Chacón, J. (2020). Modelación de la deserción universitaria mediante cadenas de Markov. doi: <http://dx.doi.org/10.15359/ru.34-1.8>
3. Guzmán Rincón, A., & Valencia Quecano, L. (2024). MODELO DE MONTE CARLO PARA LA PREDICCIÓN DE LA DESERCIÓN: HERRAMIENTA PARA LA RETROALIMENTACIÓN DE LAS POLÍTICAS PÚBLICAS EN LA EDUCACIÓN SUPERIOR. *Razón Crítica*.
4. Otero Caicedo, R., Bolívar Atuesta, S., & Palacios Caicedo, J. (2016). Análisis de la retención de estudiantes de ingeniería basado en la pérdida consecutiva de una misma asignatura. Un enfoque de Cadenas de Markov.
5. Castrillón Gómez, O., Sarache, W., & Ruiz Herrera, S. (2025). Predicción de las principales variables que conllevan al abandono estudiantil por medio de técnicas de minería de datos. *Formación Universitaria*. Obtenido de https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-50062020000600217
6. Castrillón Gómez, O., Sarache , W., & Ruiz Herrera , S. (2020). Predicción de las principales variables que conllevan al abandono estudiantil por medio de técnicas de minería de datos. *Formación Universitaria*. Obtenido de https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-50062020000600217
7. Hernández, R., Fernández, C., & Baptista, P. (2008). *Metodología de la Investigación*.
8. JIMENEZ, M. O. (2016). Análisis de los créditos y la recuperación de las cuentas por cobrar de la empresa Inmobiliaria Familia Flores S.A.C. del distrito de Villa María del Triunfo-2016. Obtenido de <https://repositorio.autonoma.edu.pe/handle/20.500.13067/691>
9. Olmedo Alonso, A., García Bernal , A., Reyes Rodríguez, A., & Tetlalmatzi Montiel, M. (2025). Cadenas de Markov absorbentes para analizar el abandono escolar de estudiantes universitarios de matemáticas. *Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=10143062>