

Índice de rendimiento estudiantil en evaluaciones en línea

María del Carmen Soto Camacho

Institución: Sección de Estudios de Posgrado e Investigación – Unidad Profesional Interdisciplinaria de Ingeniería y
Ciencia Sociales y Administrativas - IPN

Email: msotoc0600@alumno.ipn.mx

URL ORCID: <https://orcid.org/0009-0008-1007-9310>

Elena Fabiola Ruiz Ledesma

Institución: Sección de Estudios de Posgrado e Investigación – Escuela Superior de Computo - IPN

Email: eruizl@ipn.mx

URL ORCID: <https://orcid.org/0000-0002-1513-8243>

Aldo Ramírez Arellano

Institución: Sección de Estudios de Posgrado e Investigación – Unidad Profesional Interdisciplinaria de Ingeniería y
Ciencia Sociales y Administrativas - IPN

Email: aramirezar@ipn.mx

URL ORCID: <https://orcid.org/0000-0002-6782-9847>

Resumen— El desempeño académico en la educación se ve influenciado principalmente por la participación del docente. Sin embargo, en el nivel superior el rendimiento estudiantil está vinculado en mayor medida a factores cognitivos. En los últimos años han surgido nuevas herramientas tecnológicas, como las plataformas de cursos masivos abiertos en línea (MOOC, por sus siglas en inglés) y los sistemas de gestión del aprendizaje (LMS, por sus siglas en inglés), que mejoran la experiencia educativa con el acceso a cursos en línea, facilitando la gestión, el seguimiento y la evaluación del desempeño académico. Este estudio, se enfoca en las evaluaciones en línea y el análisis de los datos de interacción de los estudiantes en los exámenes en línea. Introduciendo un Índice de Rendimiento Estudiantil en Evaluaciones en Línea (IREEL), que se obtiene a partir de redes individuales y colectivas de datos de estudiantes inscritos en diversos programas de licenciatura e ingenierías del Instituto Politécnico Nacional (IPN) que presentaron evaluaciones en línea en Moodle. Los resultados muestran que el IREEL es mayor en los estudiantes que aprobaron el examen en comparación con los alumnos que no lograron aprobar.

Palabras Clave — Redes, entropía, evaluación en línea, rendimiento académico.

Abstract- Academic performance in education is mainly influenced by teacher engagement. However, at the higher education level, student performance is more strongly linked to cognitive factors. In recent years, new technological tools have emerged, such as Massive Open Online Course (MOOC) platforms and Learning Management Systems (LMS), which enhance the educational experience by providing access to online courses and facilitating course management, progress tracking, and academic performance

assessment. This study focuses on online assessments and the analysis of student interaction data in online exams. It introduces the Index of Student Performance in Online Assessments (IREEL), which is obtained from individual and collective student data networks. The study includes students enrolled in various bachelor's and engineering programs at the Instituto Politécnico Nacional (IPN), who participated in online assessments hosted on Moodle. The results show that the IREEL is higher in students who passed the exam compared to those who failed.

I. INTRODUCCIÓN

El desempeño académico en la educación superior está influenciado tanto por la participación docente como por factores cognitivos [1]. En los últimos años, la innovación en herramientas tecnológicas [2], como los Cursos Masivos Abiertos en Línea (MOOC, por sus siglas en inglés) y los Sistemas de Gestión del Aprendizaje (LMS, por sus siglas en inglés), han transformado el proceso educativo al facilitar el acceso a cursos en línea y permitir el seguimiento del progreso estudiantil [3]. A nivel global, las tasas de deserción universitaria han sido preocupantes, alcanzando un 40 % en 2018 [4]. En México, durante el ciclo 2020-2021, aproximadamente 300,000 estudiantes abandonaron sus estudios debido a bajo rendimiento académico, y en el Instituto Politécnico Nacional (IPN), la tasa de deserción en el ciclo 2022-2023 fue del 6.41 %, con una tasa de reprobación del 38.87 %. Estos datos reflejan la necesidad de mejorar la evaluación del aprendizaje para reducir el fracaso académico.

Diversos estudios han abordado el rendimiento estudiantil desde distintas perspectivas [5]. Se ha analizado el impacto de

factores socioeconómicos [6] en el éxito académico, el diseño de evaluaciones más efectivas [7] y la implementación de metodologías avanzadas para predecir el desempeño de los estudiantes [8], [9]. Además, el uso de tecnologías digitales ha permitido optimizar las evaluaciones y reducir la ansiedad relacionada con los exámenes [10]. Investigaciones recientes han empleado la minería de datos estudiantiles (EDM, por sus siglas en inglés) y técnicas de aprendizaje automático (ML, por sus siglas en inglés) para predecir tasas de éxito, fracaso y deserción [1]. Modelos basados en redes neuronales recurrentes (RNN, por sus siglas en inglés) [3], [11], unidades de memoria a corto y largo plazo (LSTM, por sus siglas en inglés) [12], [13] y redes neuronales convolucionales (CNN, por sus siglas en inglés) [14] han demostrado ser herramientas eficaces para analizar el comportamiento estudiantil y mejorar los procesos de evaluación.

A pesar de estos avances, no se ha desarrollado un índice basado en modelos de redes que represente la interacción entre estudiantes y evaluaciones. Este estudio introduce el desarrollo de un Índice de Rendimiento Estudiantil en Evaluaciones en Línea (IREEL), diseñado a partir de la entropía de redes de la interacción estudiante-evaluación en línea. A diferencia de otros enfoques, este modelo no requiere información adicional sobre variables cognitivas, factores sociales o características personales, sino que, se basa únicamente en los patrones de interacción observados en plataformas de evaluación en línea.

Este trabajo busca responder la siguiente pregunta de investigación:

¿Existe una diferencia significativa en el IREEL entre estudiantes que aprobaron y aquellos que no aprobaron?

II. METODOLOGÍA/DESARROLLO

Este estudio se llevó a cabo bajo un enfoque cuantitativo [15], centrado en la recopilación y análisis de datos numéricos para calcular el IREEL. Su objetivo es explorar la relación entre las interacciones de los estudiantes durante los exámenes en línea y su rendimiento académico. La investigación busca identificar patrones de comportamiento de los estudiantes al interactuar con el sistema de gestión de aprendizaje (LMS) durante las evaluaciones en línea para distinguir comportamientos entre los estudiantes que obtienen una calificación aprobatoria y los que no. Una calificación inferior a seis se considera reprobada, y la calificación positiva es aprobada.

La población objetivo de esta investigación incluye estudiantes de educación superior matriculados en programas escolarizados de licenciatura en el IPN. Estos estudiantes participaron en cursos que incluían exámenes en línea como parte de sus evaluaciones. La muestra se seleccionó mediante un muestreo por conveniencia, considerando la disposición de los participantes. Se recopilieron un total de 1629 exámenes de cursos como: bases de investigación clínica, fisiología humana, gestión de laboratorios, fisiopatología y computación.

A. Instrumentos y recopilación de datos

Los datos se recopilaron mediante exámenes en línea utilizando Moodle, un sistema de gestión educativa popular. El profesor titular diseñó los exámenes siguiendo las especificaciones: cada examen comprendía varios tipos de preguntas (como opción múltiple o preguntas de sí/no). Los datos se recopilaron de manera sistemática en el IPN durante los cursos semestrales entre 2018 y 2023.

En Moodle registraron datos de interacción de los usuarios que luego se descargaron en un formato estructurado de base de datos MySQL.

B. Preprocesamiento de datos

El preprocesamiento de los datos se realizó mediante procedimientos almacenados de bases de datos para extraer la siguiente información: identificador del curso, identificador del examen, identificador del usuario, identificador de cada pregunta, tiempo de respuesta de cada pregunta, calificación por pregunta, tiempo total para completar el examen y la calificación final del examen. Se omitió el uso de datos personales sensibles para proteger la confidencialidad de los estudiantes.

C. Flujo de la metodología

La secuencia de los pasos para analizar las interacciones de los estudiantes durante los exámenes en línea y crear el IREEL se describe en las siguientes secciones:

1. Construcción de la red de interacción estudiante-examen en línea.

El análisis de las trayectorias de aprendizaje a través de los sistemas LMS en entorno de aprendizaje mixto desde la perspectiva de las redes, ha permitido identificar propiedades claves tales como la emergencia, equifinalidad y la no-Aditividad [16], [17], [18].

En el presente estudio se llevó a cabo la construcción de redes individuales y colectivas a partir de los registros obtenidos de la plataforma Moodle. En las redes, cada uno de los nodos representó una de las preguntas que integraban el examen, y los enlaces entre nodos se representaron mediante las interacciones del estudiante con las preguntas durante el examen.

Para construir la red, se consideró cada evento de evaluación en orden cronológico. Por ejemplo, si un estudiante inició la evaluación con la Pregunta 1 y luego pasó a la Pregunta 2, se agregó un arco de P1 a P2. Si el estudiante volvió a la Pregunta 1 antes de responder la Pregunta 2, se agregó otro arco en sentido inverso. Este proceso se repitió hasta completar todas las interacciones de los estudiantes con el examen, como se muestra en la **Figura 1**. Las redes colectivas se

obtuvieron combinando las redes individuales y sumando los valores en la matriz de adyacencia.

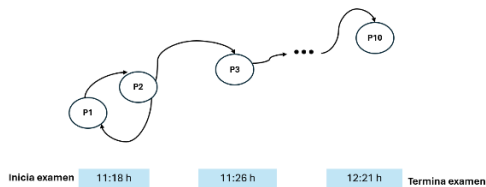


Figura 1.-Eventos ordenados cronológicamente obtenidos de la interacción alumno - examen en línea

2. Algoritmo de cobertura de cajas

Para llevar a cabo el análisis de la estructura de las redes de interacciones, se empleó el algoritmo de cobertura de cajas(box-covering) [19], el cual consiste en cubrir la red con el menor número de cajas de un determinado diámetro (\mathcal{E}).

Dada una red $G(V, E)$ obtenida de los registros de la aplicación del examen en línea, en donde V representa los nodos (preguntas) y E las conexiones (interacción entre preguntas), se construyó una red dual G' considerando los nodos de la red original G . Tomando dos nodos i y j con una distancia $\geq \mathcal{E}$ con $\mathcal{E}=2$, los nodos debían estar conectados entre sí en G' . Para ello, se elaboró una matriz en la cual se colocó la distancia mínima entre cada par de nodos en G y se determinó el diámetro de la red, el cuál debe ser el valor máximo en la matriz.

Posteriormente, se asignaron colores a los nodos, tomando en cuenta que dos nodos conectados entre sí en G' no podían ser del mismo color, buscando minimizar el número de colores. Con ello se formaron cajas que estaban integradas por los nodos del mismo color. En este caso una red dual (dos cajas) Ver Figura 2.

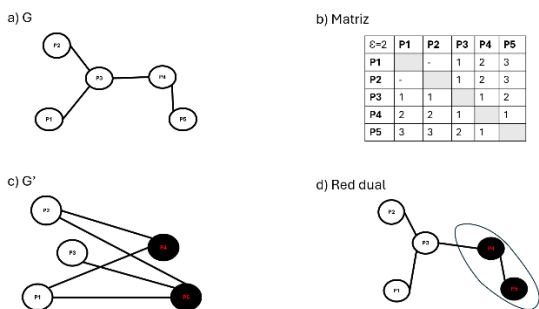


Figura 2.- Algoritmo de cubrimiento de cajas para $\mathcal{E}=2$.

Este proceso se repitió variando el diámetro(\mathcal{E}) hasta llegar al diámetro máximo $\mathcal{E} + 1$.

3. Calculo de la entropía de la red
 La entropía se utiliza para calcular el grado de complejidad dentro de un sistema [20]. En este trabajo se utilizó la entropía de Shannon para calcular la entropía en las redes de interacción.

a. Entropía de Shannon

La entropía de Shannon se define como $H(x)$ para una fuente de datos X con n eventos, en donde cada evento i tiene una probabilidad $P(x_i)$, para el cálculo de la entropía de las redes se utilizó la siguiente ecuación [21]:

$$H(x) = - \sum_{i=1}^n P(x_i) \log P(x_i) \quad (1)$$

Dónde $P(x_i)$ es la probabilidad de que ocurra el evento x_i , x_i es un elemento del evento X que puede tomar los valores $[x_1, x_2, \dots, x_n]$, y n es el número de posibles estados.

Para la red, la entropía se calculó considerando las cajas obtenidas con el algoritmo de cobertura y el número de nodos dentro de cada caja. La distribución de estas cajas permitió obtener la probabilidad y calcular la entropía en función del tamaño de la caja.

Cuando el valor de la entropía es alto, es decir cuando todos los eventos tienen la misma probabilidad de ocurrir, se alcanza la entropía máxima. Por el contrario, cuando la probabilidad de uno de los eventos se aproxima a uno y el resto tienden a cero, se tiene la entropía mínima.

4. Cálculo del IREEL

El cálculo del IREEL se basó en la medida de complejidad basada en la entropía (EMC) [22], y está definido por la siguiente ecuación:

$$IREEL = 1 - \frac{\int_a^b H(\epsilon) d\epsilon}{(b-a)(H_{\max}(a) - H_{\max}(b))} \quad (2)$$

Dónde $H(\epsilon)$ es la entropía de la red para diferentes valores de ϵ , y H_{\max} es la entropía máxima en los extremos del intervalo. Este índice mide la complejidad de las interacciones entre los estudiantes y el examen en línea, indicando una mayor complejidad cuando los valores del IREEL son más altos.

III. RESULTADOS Y DISCUSIÓN

De los datos obtenidos, se construyeron 354 redes individuales basadas en las interacciones de los estudiantes durante la aplicación de exámenes en línea. EL IREEL se calculó y se

analizó de acuerdo con la calificación obtenida ("Aprobado" para resultados de seis o más y "Reprobado" para resultados de menos de seis). A continuación, se muestra un breve resumen del cálculo del IREEL. Ver **Tabla 1**.

Tabla 1. Resumen IREEL

	<i>Mínimo</i>	<i>Máximo</i>	<i>Promedio</i>
IREEL	0.7085	0.9068	0.8371
Calificación	0	10	5.8552

Se llevó a cabo una prueba de normalidad de los valores del IREEL obtenidos tanto para los estudiantes que aprobaron como para los que no aprobaron. Los resultados en la prueba Kolmogórov-Smirnov revelan que, los datos no siguen una distribución normal, ya que los estudiantes que aprobaron ($K-S(242) = 0.931, p < 0.001$) y los que reprobaron ($K-S(111) = 0.946, p < 0.001$) presentaron un valor para la significancia menor a 0.001 para los dos grupos, se rechazó la hipótesis nula de que los datos siguen una distribución normal.

Considerando que los datos no siguen una distribución normal, fue necesario efectuar una prueba no paramétrica de Mann-Whitney U para el análisis de los valores de las medianas. Se detectó la existencia de diferencia estadísticamente significativa únicamente en el caso de las redes individuales, con $U(242, 111) = 5.3393, p < 0.001$, como lo indicó la prueba realizada. El IREEL para los estudiantes con calificación aprobatoria (mediana = 0.851) fue mayor en comparación con los estudiantes con calificación no aprobatoria (mediana = 0.842). Estos resultados evidencian una mayor complejidad en las interacciones entre los estudiantes que aprobaron la evaluación y los que no.

Por otro lado, en las redes colectivas no se encontró una diferencia significativa, como lo demuestra $U(242, 111) = 0.1947$ y $p = 0.846$. Esta falta de diferencia puede deberse a la similitud entre las dos redes colectivas de la misma evaluación construidas utilizando las interacciones de los estudiantes que aprobaron y reprobaron. El IREEL para los estudiantes con calificación aprobatoria (mediana = 0.685) fue menor en comparación con los estudiantes con calificación no aprobatoria (mediana = 0.690).

IV. CONCLUSIONES

Los resultados obtenidos indican que el IREEL es útil para evaluar la complejidad en las interacciones de los estudiantes al momento de presentar exámenes en línea y su relación con el rendimiento académico. También se observa que las interacciones individuales de los estudiantes que aprobaron son más complejas en comparación con aquellos que reprobaron, lo que sugiere un mayor compromiso o estrategias de estudio más efectivas. En contraste, las redes colectivas muestran menor complejidad y no presentan diferencias estadísticas

significativas, lo que indica que las interacciones grupales no son un factor determinante en el desempeño académico.

Al comparar estos hallazgos con estudios previos sobre modelado de trayectorias de aprendizaje a partir de datos de interacción en sistemas LMS, se han identificado patrones en estudiantes exitosos y en aquellos que no aprueban. Aunque algunos modelos basados en redes neuronales han logrado mayor precisión en la predicción del rendimiento, la ventaja de nuestro enfoque radica en su simplicidad y eficiencia, al centrarse en datos de evaluaciones de corta duración en lugar de información recopilada durante semanas o meses.

A pesar de la relevancia de estos hallazgos, se deben considerar limitaciones como la falta de diferencias significativas en las redes colectivas, lo que sugiere que otros factores, como la calidad y diversidad de las interacciones, pueden influir en los resultados. Futuros estudios profundizarán en el comportamiento de las redes colectivas para comprender mejor su impacto en el rendimiento académico.

AGRADECIMIENTOS

Agradecimientos

Este trabajo de investigación recibió financiamiento parcial a través del programa SIP con número 20240278 del Instituto Politécnico Nacional. El primer autor expresa su agradecimiento a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo económico (651b2ff07392111317337653) que respalda parcialmente este trabajo.

DECLARACIÓN ÉTICA

Para garantizar el anonimato de los estudiantes, estos fueron identificados únicamente mediante números asignados por el sistema de Moodle al registrarse como usuarios. Así mismo, a todos los participantes se les explicó en detalle el propósito y el proceso de la investigación, obteniendo su consentimiento informado de manera verbal para el uso de la información mencionada anteriormente.

REFERENCIAS

- [1] M. Chitti and M. Jayabalan, "Need for Interpretable Student Performance Prediction," presented at the IEEE 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, UK, 2020. doi: 10.1109/DeSE51703.2020.9450735.
- [2] S. Sascha, B. Maik, N. Steve, and S. Lenka, "The Cognitive-Affective-Social Theory of Learning in digital Environments (CASTLE)," *Educ Psychol Rev.* vol. 34, pp. 1–38, 2022, doi: <https://doi.org/10.1007/s10648-021-09626-5>.
- [3] J. E. Meku, B. Batchakui, R. Nkambou, and G. Okereke, "ALGORITHMS FOR THE DEVELOPMENT OF DEEP LEARNING MODELS FOR CLASSIFICATION AND PREDICTION OF BEHAVIOUR IN MOOCS," *Artificial Intelligence for Data Science in Theory and Practice*, vol. 20, pp. 41–73, 2022, doi: https://doi.org/10.1007/978-3-030-92245-0_3.
- [4] P. Prabhu, P. Valarmathie, and K. Dinakaran, "A Feature Learning-Based Model for Analyzing Students' Performance in Supportive Learning," *Intelligent Automation & Soft Computing*, vol. 36, pp. 2989–3005, 2023, doi: <https://doi.org/10.32604/iasc.2023.028659>.
- [5] R.-A. Aldo Ramirez-Arellano, B.-R. Juan, and H.-S. Luis Manuel, "Emotions, Motivation, Cognitive– Metacognitive Strategies, and

- Behavior as Predictors of Learning Performance in Blended Learning,” *Journal of Educational Computing Research*, vol. 57, no. 2, pp. 491–51, 2019, doi: <https://doi.org/10.1177/0735633117753935>.
- [6] Arya Espahbodi, Hassanali Espahbodi, Linda Espahbodi, Reza Espahbodi, Rosemary Walker, and G. Thomas White, “Environmental and university opportunity factors and CPA exam performance,” *Journal of Accounting and Public Policy*, vol. 45, pp. 0278–4254/, 2024, doi: <https://doi.org/10.1016/j.jaccpubpol.2024.107208>.
- [7] Fernandez Moya, Maria, Cuadros, Pedro, and Salvador, Carlos, “HOMOGENIZATION OF WORLD UNIVERSITY EDUCATION: A PROPOSAL BASED ON EXAMS,” presented at the 14TH INTERNATIONAL TECHNOLOGY, EDUCATION AND DEVELOPMENT CONFERENCE, Valencia, SPAIN: ATED-INT ASSOC TECHNOLOGY EDUCATION & DEVELOPMENT, 2020, pp. 3889–3892. doi: 10.21125/inted.2020.1084.
- [8] S. Xu, “The Application of Multimedia and Deep Learning in the Integration of Professional and Innovative Education in Colleges,” *International Journal of Information Technologies and Systems Approach*, vol. 16, no. 2, pp. 1–13, 2023.
- [9] H. Huang, S. Yuan, T. He, and R. Hou, “Use of Behavior Dynamics to Improve Early Detection of At-risk Students in Online Courses,” *Mobile Netw Appl*, vol. 27, pp. 441–452, 2022, doi: <https://doi.org/10.1007/s11036-021-01844-z>.
- [10] Reem AlSufayan and Dina Abdel Salam, “Achievement Emotions in Paper-Based Exams vs. Computer-Based Exams: The Case of a Private Saudi University,” *International Journal of Online Pedagogy and Course Design (IJOPCD)*, vol. 13, no. 1, p. 21, 2023, doi: 10.4018/IJOPCD.322084.
- [11] Y. He, R. Chen, X. Li, C. Hao, S. Liu, and G. Zhang, “Online At-Risk Student Identification Using RNN-GRU Joint Neural Networks,” *Journal Information*, vol. 1, pp. 1–12, 2020, doi: <https://doi.org/10.3390/info11100474>.
- [12] M. A. Salam, M. A. El-Fatah, and N. F. Hassan, “Automatic grading for Arabic short answer questions using optimized deep learning model,” *PLoS ONE*, vol. 17, no. 8, pp. 1–41, 2022, doi: <https://doi.org/10.1371/journal.pone.0272269>.
- [13] S. Qu, K. Li, B. Wu, S. Zhang, and Y. Wang, “Predicting Student Achievement Based on Temporal Learning Behavior in MOOCs,” *appls*, vol. 1, pp. 924–941, 2019, doi: <https://doi.org/10.3390/app9245539>.
- [14] Z. Hancox and S. D. Relton, “Temporal Graph-Based CNNs (TG-CNNs) for Online Course Dropout Prediction,” *Lecture Notes in Computer Science*, vol. 13515, pp. 357–367, 2022, doi: https://doi.org/10.1007/978-3-031-16564-1_34.
- [15] R. H. Sampieri and C. P. Mendoza Torres, *Metodología de la investigación Las rutas cuantitativa, cualitativa y mixta*. Mc Graw Hill Education, 2018.
- [16] Ortiz-Vilchis, Pilar and Aldo Ramirez-Arellano, “Learning Pathways and Students Performance: A Dynamic Complex System,” *Entropy*, vol. 25, no. 291, pp. 1–14, 2023, doi: <https://doi.org/10.3390/e25020291>.
- [17] A. Ramirez, “Students learning pathways in higher blended education: An analysis of complex networks perspective,” *Elsevier Ltd*, vol. 141, pp. 1–15, 2019, doi: <https://doi.org/10.1016/j.compedu.2019.103634>.
- [18] A. Ramirez-Arellano, J. M. S. Almira, and J. B. Reyes, “Fractional Online Learning Rate: Influence of Psychological Factors on Learning Acquisition,” *Nonlinear Dynamics, Psychology, and Life Sciences*, vol. 26, pp. 289–313, 2022.
- [19] Chaoming Song, Lazaros K Gallos, Shlomo Havlin, and Hernán A Makse, “How to calculate the fractal dimension of a complex network: the box covering algorithm,” *Journal of Statistical Mechanics: An IOP and SISSA journal Theory and Experiment*, pp. 1–16, 2007, doi: 10.1088/1742-5468/2007/03/P03006.
- [20] Berthold Bein MD, “Entropy,” *Best Practice & Research Clinical Anaesthesiology*, vol. 20, no. 1, pp. 101–109, 2006, doi: <https://doi.org/10.1016/j.bpa.2005.07.009>.
- [21] Shannon, Claude E, “A mathematical theory of communication,” *Bell system technical journal*, vol. 27, pp. 379–423, 1948.
- [22] Pilar Ortiz-Vilchis and Aldo Ramirez-Arellano, “An Entropy-Based Measure of Complexity: An Application in Lung-Damage,” *Entropy*, vol. 24, pp. 1–13, 2022, doi: <https://doi.org/10.3390/e24081119>.