

Personalización del Aprendizaje mediante Ciencia de Datos:
Estrategias y Aplicaciones en la Educación Superior

Personalized Learning through Data Science: Strategies and
Applications in Higher Education

Lesdiel Antonio Pinzón Aparicio¹

¹Universidad de Panamá, pinzón@up.ac.pa, <https://orcid.org/0000-0002-4475-2015>, Panamá

Información del Artículo

Trazabilidad:

Recibido 10-12-2024

Revisado 11-12-2024

Aceptado 26-12-2024

Palabras Clave:

Ciencia de datos
Personalización del aprendizaje
Educación superior
Analíticas de aprendizaje

Keywords:

Data science
Personalization of learning
Higher education
Learning Analytics

RESUMEN

Este estudio analiza el impacto de la ciencia de datos en la educación, destacando su capacidad para personalizar las experiencias de aprendizaje, se orienta a explorar el impacto de las herramientas como la minería de datos, el aprendizaje automático y las analíticas de aprendizaje en las necesidades individuales, anticipar desafíos y optimizar procesos educativos en un marco ético que garantice la privacidad. Se adopta un enfoque documental, recopilando y analizando literatura científica y técnica publicada entre 2015 y 2023. Las fuentes incluyen bases de datos como Scopus, Google Scholar y ResearchGate. Los hallazgos fundamentan que el uso de dashboards personalizados y sistemas de tutoría inteligente mejoran el compromiso estudiantil y la efectividad docente. Además, herramientas predictivas permiten identificar patrones de deserción con alta precisión, diseñando políticas proactivas que fomentan la retención. Se concluye que la ciencia de datos representa una herramienta transformadora para personalizar la educación superior, optimiza el aprendizaje y promueve una planificación más eficiente. A pesar de los retos éticos, su implementación adecuada puede impulsar una educación más inclusiva y efectiva, equilibrando el avance tecnológico con la protección de los derechos estudiantiles.

ABSTRACT

This study analyzes the impact of data science in education, highlighting its ability to personalize learning experiences, aims to explore the impact of tools such as data mining, machine learning and learning analytics on individual needs, anticipate challenges and optimize educational processes within an ethical framework that ensures privacy. A documentary approach is adopted, collecting and analyzing scientific and technical literature published between 2015 and 2023. Sources include databases such as Scopus, Google Scholar and ResearchGate. The findings support the idea that the use of personalized dashboards and intelligent tutoring systems improve student engagement and teaching effectiveness. In addition, predictive tools allow identifying attrition patterns with high accuracy, designing proactive policies that encourage retention. It is concluded that data science represents a transformative tool for personalizing higher education, optimizing learning and promoting more efficient planning. Despite ethical challenges, its proper implementation can drive more inclusive and effective education, balancing technological advancement with the protection of student rights.

INTRODUCCIÓN

La ciencia de datos en la personalización del aprendizaje se orienta a mejorar la eficacia del aprendizaje, incrementar el compromiso estudiantil, y abordar las necesidades específicas de cada alumno para optimizar su rendimiento. Se parte desde la base de que los estilos de aprendizaje son elementos cruciales en el crecimiento humano, por lo que la educación aspira a potenciar las habilidades considerando la variedad de saberes en los individuos, que fomente un discernimiento más eficaz y enriquecedor (Escobar et al., 2022). El estudio denominado impacto de las analíticas de aprendizaje en discusiones asincrónicas en educación superior de Cerro et al. (2020) investigó cómo las analíticas de aprendizaje pueden mejorar el seguimiento

y evaluación en entornos de aprendizaje colaborativo en línea. El hallazgo principal fue que el uso de métricas específicas favorece la personalización del aprendizaje y mejora la práctica docente, destacando el potencial de herramientas como DIANA para optimizar los resultados en diferentes contextos educativos. Concluye que estas soluciones tecnológicas pueden impulsar significativamente la calidad educativa y la toma de decisiones basada en datos.

El trabajo exploró de Vázquez et al. (2021) cómo los dashboards personalizados ayudan a monitorizar el desempeño estudiantil en entornos educativos dinámicos. El estudio concluyó que el diseño modular de herramientas de visualización es clave para personalizar la experiencia educativa, maximizando la efectividad de la enseñanza y alineando los objetivos educativos con las necesidades individuales de los estudiantes.

La Red Española de Analíticas de Aprendizaje (SNOLA) identificó desafíos éticos en el uso de datos estudiantiles, destacando la importancia de garantizar la privacidad y el manejo responsable de la información. El estudio subraya la necesidad de marcos éticos sólidos para equilibrar el potencial de estas tecnologías con las preocupaciones éticas y legales inherentes (Martínez et al., 2020).

Hoy en día, la Data Science es ampliamente empleada y con una tendencia ascendente en la obtención de conocimiento e información de grandes cantidades de datos, creando varias oportunidades para la creación de aplicaciones en diversos sectores. Un sector crucial en nuestro país es el educativo. Por lo que, según Orihuela Maita (2019), es crucial anticiparse a los eventos futuros, como por ejemplo, que en una carrera universitaria la mitad de los alumnos repiten un curso específico y así sucesivamente.

Por su parte, Khalil & Ebner (2015) en su modelo destacaron cómo integrar big data, minería de datos y aprendizaje automático para analizar el comportamiento estudiantil. El estudio concluyó que las instituciones pueden mejorar la personalización del aprendizaje y optimizar recursos institucionales mediante la implementación estratégica de analíticas de datos anónimos y actividades de aprendizaje específicas.

Con la intención de explorar y desarrollar un modelo predictivo de deserción enfocado en estudiantes universitarios no tradicionales, Huo et al. (2023) proponen un modelo basado en aprendizaje automático, utilizando características como el rendimiento académico y la interacción con recursos universitarios, por medio del cual se logró una alta precisión al identificar estudiantes en riesgo de abandonar. Concluyeron que las herramientas basadas en ciencia de datos no solo son efectivas para identificar patrones de deserción, sino que también permiten diseñar políticas proactivas y recomiendan que las universidades adopten enfoques integrales que combinen datos históricos con análisis contextual para prevenir la deserción de manera eficiente.

Por su parte, Márquez-Vera et al. (2016) identifican estudiantes de secundaria con riesgo de abandono escolar temprano mediante técnicas de minería de datos, contribuyendo a prevenir este fenómeno con intervenciones adecuadas. Los modelos predictivos aplicados en el estudio, basados en algoritmos como Árboles de Decisión y Random Forest, alcanzaron una precisión superior al 80%. Concluyeron que las tecnologías de minería de datos tienen un gran potencial para ser integradas en sistemas educativos, permitiendo identificar riesgos de manera temprana. Resaltan la importancia de la colaboración entre educadores, psicólogos y analistas de datos para interpretar y actuar sobre los resultados, mejorando las tasas de retención en secundaria.

Este estudio se orienta a explorar cómo la integración de la ciencia de datos, la minería de datos y el aprendizaje automático puede transformar los procesos educativos en entornos de educación superior, centrándose en la personalización de experiencias de aprendizaje. Se analizan y sintetizan estrategias basadas en datos que permitan identificar las necesidades individuales de los estudiantes, anticipar desafíos como la deserción escolar y fomentar intervenciones proactivas para maximizar el rendimiento académico, el compromiso estudiantil y la efectividad docente, todo ello dentro de un marco ético que garantice la privacidad y el manejo responsable de la información.

MATERIALES Y MÉTODOS

Este estudio se aborda desde un enfoque documental, el cual, según Reyes & Carmona (2020), tiene como finalidad orientar la investigación desde dos perspectivas, inicialmente, vinculando datos ya existentes provenientes de diversas fuentes y luego ofreciendo una perspectiva panorámica y sistemática de un tema específico desarrollado en varias fuentes dispersas.

Se basa en la recopilación, análisis y síntesis de información proveniente de fuentes académicas, científicas y técnicas relacionadas con la personalización del aprendizaje mediante ciencia de datos. Este método permite explorar el estado del arte en la materia, identificar tendencias, aplicaciones prácticas y desafíos éticos asociados al tema (Guevara, 2016).

Las fuentes de información son todos los recursos académicos de acceso pago y de libre acceso, empleando bases de datos como Scopus, Google Scholar y ResearchGate, considerando estudios publicados entre 2015 y 2023. Además, se tomaron en cuenta informes institucionales.

Se utilizaron las siguientes combinaciones de palabras clave: “ciencia de datos en educación”, “analíticas de aprendizaje” y “personalización educativa”, para orientar la búsqueda en bases de datos académicas. La selección documental se realizó siguiendo criterios de:

- Relevancia del tema de personalización educativa.
- Recencia de las publicaciones, priorizando los últimos 8 años.
- Impacto de los estudios en el campo de las analíticas de aprendizaje.

Esta metodología facilita la creación de una visión holística sobre cómo la ciencia de datos revoluciona el aprendizaje personalizado, estableciendo un fundamento sólido para futuras investigaciones y aplicaciones prácticas en el ámbito de la educación superior.

RESULTADOS Y DISCUSIÓN

A continuación, se muestra una serie de aportaciones a la temática que permiten un mejor acercamiento a la personalización del aprendizaje mediante ciencia de datos.

Tabla 1: Ciencia de datos y personalización del aprendizaje

Autor-año	Objetivo	Metodología	Hallazgo	Conclusiones
Bernal & Prados (2024).	Analizar la producción científica sobre los sistemas de tutoría inteligente en los procesos de enseñanza-aprendizaje.	Criterios PRISMA.	Existe una alta producción científica y mejoras en el rendimiento académico de los estudiantes gracias a los STI y la importancia de una coordinación interdisciplinar.	Los Sistemas de Tutoría Inteligente influyen positivamente en los procesos de enseñanza-aprendizaje al mejorar la participación y el compromiso de los estudiantes.
Teran-Pazmiño et al. (2024).	Determinar los argumentos de mayor relevancia respecto al uso de la IA en la personalización del aprendizaje.	Teórica- Documental	El uso de la tecnología beneficia la personalización del aprendizaje debido a que mediante las herramientas digitales se mejora y atiende las necesidades educativas de los alumnos.	El uso de aplicaciones y herramientas digitales permite a los docentes y estudiantes acercarse a los contenidos de una mejor manera.
Cerro Martínez et al. (2020).	Evaluar cómo las analíticas de aprendizaje mejoran el seguimiento y evaluación en discusiones asincrónicas.	Aplicación de métricas en plataformas colaborativas usando la herramienta DIANA.	Las métricas específicas optimizan la personalización y mejoran la práctica docente.	Las herramientas basadas en datos pueden transformar la calidad educativa y la toma de decisiones.
Vázquez et al. (2021).	Explorar el uso de dashboards personalizados para monitorizar el desempeño estudiantil.	Diseño modular de herramientas de visualización aplicadas en entornos dinámicos.	Los dashboards mejoran la efectividad educativa al alinear objetivos con necesidades individuales.	Es prioritario la adaptación de las herramientas tecnológicas para maximizar el aprendizaje personalizado.
Martínez Monés et al. (2020).	Analizar los desafíos éticos de las analíticas de	Revisión bibliográfica.	Existe una falta de marcos éticos sólidos para el manejo de datos educativos.	Es fundamental equilibrar el potencial de estas tecnologías

	aprendizaje en educación superior.			con la privacidad y la ética.
Huo et al. (2023)	Desarrollar un modelo predictivo para la deserción universitaria utilizando aprendizaje automático.	Análisis de datos académicos e interacciones estudiantiles mediante machine learning.	Se logró identificar estudiantes en riesgo con alta precisión.	La ciencia de datos permite diseñar políticas educativas proactivas y efectivas.
Márquez-Vera et al. (2016).	Identificar estudiantes en riesgo de abandono escolar temprano.	Minería de datos con algoritmos como Árboles de Decisión y Random Forest.	Los modelos predictivos alcanzaron una precisión superior al 80%.	La colaboración interdisciplinaria mejora la interpretación de los resultados y las tasas de retención.
Orihuela Maita (2019).	Anticipar eventos académicos como tasas de repetición en carreras universitarias.	Análisis predictivo basado en tendencias educativas nacionales.	La ciencia de datos ayuda a identificar y abordar desafíos educativos recurrentes.	La planificación basada en datos mejora la eficiencia educativa.

La figura 2 permite visualizar la relación entre ciencia de datos y personalización del aprendizaje, enfocando el aporte en las herramientas de análisis que identifican patrones únicos de aprendizaje, posibilitan predicciones para intervenciones tempranas y en donde los principales beneficios son el incremento en el rendimiento estudiantil, mayor compromiso y retención. A continuación se describen algunos elementos presentes en esta relación, basándonos en el aporte de Akyuz (2020), en su estudio sobre los efectos de los sistemas de tutoría inteligente en el aprendizaje personalizado.

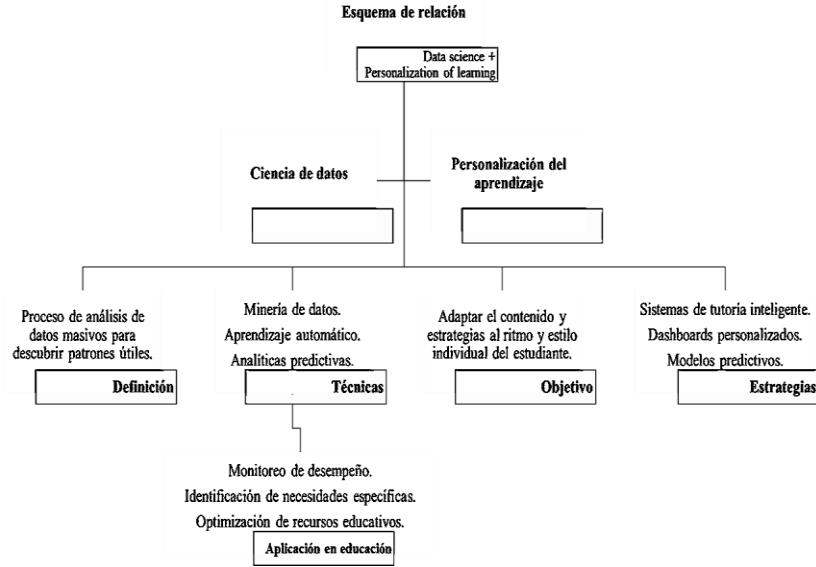


Fig. 1. Ciencia de datos y personalización del aprendizaje

El avance en la integración de tecnologías emergentes en el ámbito educativo, como los sistemas de tutoría inteligente y las analíticas de aprendizaje, ha transformado las dinámicas de enseñanza y aprendizaje, logrando un impacto significativo en la personalización y optimización de procesos educativos (Lescano et al., 2024). Con base en los resultados que se muestran en la tabla 1, las herramientas digitales personalizan la enseñanza al adaptarse a las necesidades individuales, se sustenta en los aportes presentados que la alineación de herramientas tecnológicas con objetivos específicos optimiza la efectividad educativa, promoviendo un aprendizaje centrado en el estudiante.

Sin embargo, la implementación de estas tecnologías trae consigo desafíos éticos y prácticos, por lo que se requiere el desarrollo de marcos éticos sólidos para garantizar la privacidad y el uso responsable de los datos educativos, una preocupación que también es relevante al diseñar políticas basadas en predicciones. Como bien señalan Cabero & Martínez (2019), las tareas desarrolladas por el estudiante de forma individual o colectiva en un entorno digital están destinadas a obtener un aprendizaje específico, por lo que se trata de aprovechar el potencial tecnológico y proteger los derechos de los estudiantes, lo que también ha sido señalado como esencial en la literatura reciente sobre analíticas de aprendizaje.

Finalmente, la utilidad de la ciencia de datos en la educación queda ampliamente respaldada, se ha evidenciado que el análisis predictivo permite anticipar desafíos educativos recurrentes, promoviendo una planificación más eficiente (Norambuena et al., 2022). A pesar de las preocupaciones éticas, el consenso general sugiere que estas herramientas, cuando se implementan adecuadamente, tienen el potencial de transformar la educación. Esto requiere no solo avances tecnológicos, sino también un enfoque integral que considere aspectos pedagógicos, éticos y contextuales. Así, la integración de tecnologías avanzadas en educación no solo personaliza el aprendizaje, sino que también permite a los docentes tomar decisiones más informadas y basadas en datos, marcando un camino hacia una educación más inclusiva y efectiva.

CONCLUSIÓN

El estudio fundamenta que la integración de la ciencia de datos en la educación superior se traduce en una transformación significativa hacia la personalización del aprendizaje. A través de técnicas como el análisis predictivo, la minería de datos y el aprendizaje automático, es posible anticipar desafíos educativos como la deserción estudiantil y personalizar las experiencias educativas para maximizar el rendimiento académico, la retención y el compromiso de los estudiantes. Los hallazgos destacan el papel crucial de herramientas como dashboards personalizados y sistemas de tutoría inteligente, que alinean las necesidades individuales de los estudiantes con los objetivos educativos, optimizando la práctica docente y mejorando la calidad de la enseñanza. A pesar de que el uso de estas tecnologías plantea desafíos éticos, en particular la privacidad y el uso responsable de los datos estudiantiles, el equilibrio entre tecnología y ética es esencial para garantizar un impacto positivo y sostenible en la educación superior. En este sentido, el estudio proporciona una base sólida para futuras investigaciones y aplicaciones prácticas que continúen explorando el potencial de la ciencia de datos en la creación de entornos educativos más inclusivos y efectivos.

REFERENCIAS

- Akyuz, Y. (2020). Effects of Intelligent Tutoring Systems (ITS) on Personalized Learning (PL). *Creative Education*, 11(6), Article 6. <https://doi.org/10.4236/ce.2020.116069>
- Bernal, N. C., & Prados, M. Á. H. (2024). Impacto de los sistemas de tutoría inteligente. Una revisión sistemática. *EduTec, Revista Electrónica de Tecnología Educativa*, 89, Article 89. <https://doi.org/10.21556/edutec.2024.89.3025>
- Cabero, J., & Martínez, A. (2019). *Las tecnologías de la información y comunicación y la formación inicial de los docentes: Modelos y competencias digitales*. <https://doi.org/10.30827/profesorado.v23i3.9421>
- Cerro Martínez, J. P., Guitert Catasús, M., & Romeu Fontanillas, T. (2020). Impact of using learning analytics in asynchronous online discussions in higher education. *International Journal of Educational Technology in Higher Education*, 17(1), 39. <https://doi.org/10.1186/s41239-020-00217-y>
- Escobar, B. R. P., Salazar, C. A. H., Samekash, M. L. W., & Medina, J. L. R. (2022). Estilos de aprendizaje y rendimiento académico en el área de comunicación con enfoque de sistemas. *Revista de Ciencias Sociales (Ve)*, 28(5), 48-62. <https://www.redalyc.org/journal/280/28071845004/html/>
- Guevara Patiño, R. (2016). El estado del arte en la investigación: ¿análisis de los conocimientos acumulados o indagación por nuevos sentidos? *Revista Folios*, 1(44), 165-179. <https://doi.org/10.17227/01234870.44folios165.179>
- Huo, H., Cui, J., Hein, S., Padgett, Z., Ossolinski, M., Raim, R., & Zhang, J. (2023). Predicting Dropout for Nontraditional Undergraduate Students: A Machine Learning Approach. *Journal of College Student Retention: Research, Theory & Practice*, 24(4), 1054-1077. <https://doi.org/10.1177/1521025120963821>
- Khalil, M., & Ebner, M. (2015). *Learning Analytics: Principles and Constraints*. 1789-1799. <https://www.learntechlib.org/primary/p/151455/>
- Lescano-Veloz, A., Amaiquema-Gil, S., Reigosa-Lara, A., & Tobar-Farias, G. (2024). Integración de Tecnologías Digitales Emergentes para Mejorar el Proceso de Enseñanza-Aprendizaje en la

- Asignatura de Robótica en la Formación Tecnológica. *MQRInvestigar*, 8, 247-274.
<https://doi.org/10.56048/MQR20225.8.4.2024.247-274>
- Márquez-Vera, C., Cano, A., Romero, C., Noaman, A. Y. M., Mousa Fardoun, H., & Ventura, S. (2016). Early dropout prediction using data mining: A case study with high school students. *Expert Systems*, 33(1), 107-124. <https://doi.org/10.1111/exsy.12135>
- Martínez Monés, A., Dimitriadis Damoulis, I., Acquila Natale, E., Álvarez, A., Caeiro Rodríguez, M., Cobos Pérez, R., Conde González, M. Á., García Peñalvo, F. J., Hernández Leo, D., Menchaca Sierra, I., Muñoz Merino, P. J., Ros, S., & Sancho Vinuesa, T. (2020). Achievements and challenges in learning analytics in Spain: The view of SNOLA. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), 187. <https://doi.org/10.5944/ried.23.2.26541>
- Norambuena, J. M., Badilla-Quintana, M. G., & Angulo, Y. L. (2022). Modelos predictivos basados en uso de analíticas de aprendizaje en educación superior: Una revisión sistemática. *Texto libre*, 15. <https://www.redalyc.org/journal/5771/577170677011/html/>
- Orihuela Maita, G. Y. (2019). *Aplicación de Data Science para la predicción del rendimiento académico de los estudiantes de la Facultad de Ingeniería de Sistemas de la Universidad Nacional del Centro del Perú* [Tesis de Pregrado, Universidad Nacional del Centro del Perú]. <http://repositorio.uncp.edu.pe/handle/20.500.12894/5837>
- Reyes, L., & Carmona, F. A. (2020). *La investigación documental para la comprensión ontológica del objeto de estudio*. <https://bonga.unisimon.edu.co/handle/20.500.12442/6630>
- Teran-Pazmiño, E., Cadena-Morales, L., González-González, Lady, Guamán-Sánchez, N., & León-Flores, M. (2024). Tecnología y Personalización del Aprendizaje. *Revista Científica Retos de la Ciencia*, 1, 115-129. <https://doi.org/10.53877/rc.8.19e.202409.10>
- Vázquez-Ingelmo, A., García-Peñalvo, F. J., & Therón, R. (2021). Towards a Technological Ecosystem to Provide Information Dashboards as a Service: A Dynamic Proposal for Supplying Dashboards Adapted to Specific Scenarios. *Applied Sciences*, 11(7), Article 7. <https://doi.org/10.3390/app11073249>