



CIENCIA DE DATOS ESTUDIANTILES, TENDENCIAS HACIA UNA EDUCACIÓN INTELIGENTE

María del Carmen Soto Camacho

Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas

Instituto Politécnico Nacional

msotoc0600@alumno.ipn.mx

<https://orcid.org/0009-0000-9692-0773>

Claudia Marina Vicario Solórzano

Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas

Instituto Politécnico Nacional

cvicario@ipn.mx

<https://orcid.org/0000-0003-0144-3607>

Abstract

La ciencia de datos ha revelado tener aplicación en una variedad de ámbitos, incluyendo el área de la educación (Aljawarneh & Lara, 2021), con un enfoque de la ciencia de datos encaminado a ampliar el desarrollo de los datos educativos existentes y futuros (Husfeldt, Schiller, & Alt, 2020). Este artículo tiene como objetivo realizar una revisión acerca de la ciencia de datos educativos, así como describir y analizar las tendencias de este foco de estudios de investigación, que buscan brindar aportaciones encaminadas hacia la oferta de una educación de calidad.

Data Science has been found to have applications in a variety of areas, one of them being education (Aljawarneh & Lara, 2021), as a data science approach aimed at expanding the development of existing and future educational data (Husfeldt, Schiller, & Alt, 2020). This article aims to conduct review of educational data science, as well as describe and analyze the trends of this focus of research approach, which could provide contributions to offer quality education.

Palabras clave:

Educational data science, Big Data, e-learning, LMS, Success rate, Recommender system, educational performance, quality education.

Los avances en sistemas de enseñanza no eran muy alentadores antes de la pandemia, sin embargo, con la crisis sanitaria por COVID-19

se hizo evidente un aplazamiento en cuatro de cada cinco países en materia de enseñanza de calidad (Naciones Unidas, 2023). Para



alcanzar el objetivo número 4 de desarrollo sostenible (Educación de calidad) de la UNESCO, resulta inevitable generar mejoras en la infraestructura escolar con un enfoque hacia la transformación digital y la adopción de nuevas tendencias.

Los bajos niveles de acceso a tecnologías de la información y las comunicaciones (TIC) son un claro ejemplo de las limitantes en el uso de datos e información para enriquecer los sistemas de educación y enseñanza (Naciones Unidas, 2023). Entre las tendencias de reciente nacimiento, se puede contemplar a la ciencia de datos, la cual en términos generales obedece a un gran número de algoritmos, metodologías y tecnologías que recopilan datos, los integran y analizan a fin de crear modelos, patrones y predicciones.

A lo largo de este artículo se aborda de forma accesible y general la forma en que se aplica la ciencia de datos educativos.

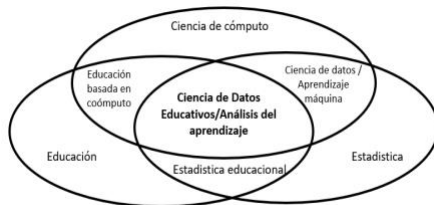


Figura 1. Áreas relacionadas con la ciencia de Datos Educativos

Ciencia de datos educativos

En los últimos años se ha presentado un importante incremento en la generación de datos a un ritmo sorprendente (Poulova, Klimova, & Mikulecká, 2019), esto como consecuencia del constante uso de dispositivos móviles, el desarrollo de aplicaciones y la implementación de sensores a un sinfín de instrumentos que utilizamos en el día a día a causa de la aparición del internet de las cosas (IoT). Este incremento en la generación de

datos se ve reflejado en el creciente desarrollo de recursos de aprendizaje electrónico, software educativo y en general al uso de internet como una herramienta de apoyo para la formación de los estudiantes (Romero & Ventura, 2020), el cual da como resultado el hecho de orientar el análisis de datos hacia la educación.

La aplicación de la ciencia de datos educativos en distintos estudios, por lo general se complementa con teorías pedagógicas, herramientas informáticas y la aplicación de métodos estadísticos, para procesar los datos masivos que se recopilan durante los cursos de enseñanza, dichos estudios pueden ofrecer modelos más eficientes de gestión del sistema de enseñanza, mejoras en la calidad de la educación, y el aprendizaje inteligente (Dong, 2019). “Mientras más datos podamos combinar y utilizar de forma significativa, más podemos decir acerca del aprendizaje y la calidad del sistema educativo” (Husfeldt, Schiller, & Alt, 2020).

Aplicaciones de la ciencia de datos educativos en la actualidad

Como ya previamente se mencionó, la ciencia de datos educativos nace como consecuencia de la alta generación de datos educativos, dentro de los cuales se pueden incluir aquellos datos relativos a los estudiantes, el comportamiento del aprendizaje, algunos factores psicológicos y de cognición (Xiaomei, Fuli, Jinzhou, & Teng, 2021).

La ciencia de datos educativos en la actualidad tiene aplicación en la gestión educativa, métodos innovadores para la docencia, modelos y aportes de investigación. Algunos ejemplos remarcables de estos casos de aplicación por mencionar algunos son las predicciones del rendimiento académico de los estudiantes, las sugerencias de cursos y



materiales de aprendizaje, la recomendación a postularse a empleos que coinciden con el perfil académico del estudiante, etc.

A través de múltiples estudios se ha revelado gracias a la ciencia de datos educativos que es posible generar predicciones sobre el desempeño de los estudiantes en las asignaturas que cursarán en los próximos periodos escolares, a este tipo de estudios se le denomina investigación predictiva (Xiaomei, Fuli, Jinzhou, & Teng, 2021).

Las tasas de deserción, retención y atracción estudiantil son un claro ejemplo del tipo de indicadores en los que las instituciones educativas, e incluso las autoridades de gobierno toman decisiones importantes, apoyadas en el análisis de datos. Como es bien sabido la deserción de estudiantes de nivel superior en América Latina presentó un considerable aumento a raíz de la crisis por COVID-19.

Los altos índices de deserción pueden estar asociadas a diferentes factores, por medio de la ciencia de datos es posible clasificar esos factores, se puede predecir con exactitud en qué medida cada uno de estos factores impacta. En los sistemas de enseñanza a través de la correcta interpretación de los datos es posible entender y predecir el comportamiento de los estudiantes y los factores que a su alrededor impactan el abandono escolar (Kuz & Morales, 2023), con ello es posible generar estrategias que permitan disminuir los altos índices de deserción.

Otro importante aporte se centra en detectar señales de alarma con respecto al comportamiento de los estudiantes, al registrar y analizar de manera constante el comportamiento de los estudiantes es posible detectar conductas inusuales, con ello se ha logrado brindar apoyo a estudiantes en los que se identificaron señales de depresión o con

conductas inadecuadas, para ello se han desarrollado sistemas de decisiones de advertencia sobre cambios en el comportamiento de estudiantes universitarios, gracias a la ciencia de datos (Xiaomei, Fuli, Jinzhou, & Teng, 2021).

Como aplicación de la ciencia de datos educativos, también podemos mencionar el aporte que ésta ha brindado para adoptar una visión de reformar no solo algunos planes de estudio, sino que el entorno en el que se imparte la educación también influye en el desempeño del aprendizaje de los estudiantes, y algunos estudios mencionan la importancia de construir ecosistemas de aprendizaje flexibles y abiertos (Casarosa, Ruggieri, Salvatori, Simi, & Turbanti, 2020).

Otros estudios se centran en el análisis del uso de plataformas y aplicaciones de cursos en línea para el apoyo en el proceso de enseñanza, mediante la ciencia de datos y algoritmos de aprendizaje automático se busca identificar si este tipo de sistemas influyen en el proceso de aprendizaje de los estudiantes, ya sea de forma positiva o negativa, es decir, si los estudiantes manifiestan motivación o compromiso al realizar actividades en estos sistemas y aplicaciones (Salas-Rueda, Salas-Rueda, Salas-Rueda, & Salas-Silis, 2018).

La ciencia de datos en los nuevos espacios educativos

Cada vez son más los casos en que tanto instituciones académicas y profesores utilizan los sistemas de gestión del aprendizaje (LMS), Estos sistemas tienen la particularidad de facilitar y promover el proceso de enseñanza – aprendizaje, fomentan la constante participación de los estudiantes antes, durante y después del tiempo destinado a impartir clases. Estos sistemas son un claro ejemplo del



auge de la ciencia de datos estudiantiles mediante técnicas de aprendizaje automático.

Los LMS llevan un registro puntual de la interacción de los estudiantes en el aula de clases, de las actividades realizadas, incluso de su participación. Un estudio en el que se contó con la participación de profesores de la UNAM (Salas Rueda, Eslava Cervantes, & Prieto Larios, 2020) arrojó, que la adopción de LMS como Moodle impactan de forma positiva la participación y comunicación en el proceso de aprendizaje no solo para los estudiantes, los profesores también reportan mejoras en la evaluación de exámenes y el fomento a la participación en debates online.

Con los datos obtenidos de este tipo de estudios se obtiene evidencia de que los docentes pueden crear, organizar y actualizar los modelos de enseñanza en relación con los datos que los sistemas LMS van arrojando del análisis de la interacción de los estudiantes.

Otro ejemplo de entornos educativos en los que la ciencia de datos hace presencia es en los video juegos enfocados a la educación, los cuales se emplean en diferentes áreas del aprendizaje tanto para la comprensión de conceptos, para simular experimentos científicos o fenómenos de distintos tipos (Canan, Aysun, & Tülin, 2022), en algunos casos al finalizar la experiencia del videojuego se aplican algunos cuestionarios para recolectar datos, los cuales posteriormente son evaluados para realizar análisis del contenido, los cuales resultan útiles para realizar mejoras en las siguientes versiones de estos productos.

Algunos de estos videojuegos se diseñaron y desarrollaron como herramientas para la recolección de datos, como lo son observaciones acerca de la iteración del estudiante, que les permiten a los docentes

evaluar el comportamiento del estudiante ante ciertas circunstancias.

Conclusiones

El nivel de calidad en la educación es importante en cualquier nación, tanto que de hecho se maneja como uno de los principales objetivos de desarrollo, por lo tanto, existe una amplia variedad de trabajos de investigación en torno a responder un sinnúmero de interrogantes tales como ¿Qué factores son relevantes para lograr un sistema de enseñanza de calidad?, ¿Cómo se alcanza una educación de calidad? Y ¿Cómo medir el progreso de la educación de una sociedad? Todos los anteriores son ejemplos de estudios destacados que se han llevado a cabo mediante técnicas y metodologías de ciencia de datos, los profesores hacen uso de la tecnología y los nuevos ambientes de educación para crear

Referencias

- Husfeldt, V., Schiller, D., & Alt, S. (2020). The expanded potential of educational data. *Wissenschaft & Praxis*, 298-302.
- Aljawarneh, S., & Lara, J. A. (2021). Data science for analyzing and improving educational processes. *J Comput High Educ*, 545-550.
- Canan, L.-Ş., Aysun, Ö.-K., & Tülin, S. (2022). Educational Games in Science Center: Experiences of Pre-Service Science Teachers. *SAGE Open Journals*.
- Casarosa, V., Ruggieri, S., Salvatori, E., Simi, M., & Turbanti, S. (2020). Educational ecosystems for Information Science: the case of the University of Pisa. *EDUCATION FOR INFORMATION*, 119-138.
- Dong, J. (2019). Analysis and Designing of Educational Big Data Mining System for Higher Education. *8th International Conference on Social Science and Education Research (SSER)*. CHINA.



Kuz, A., & Morales, R. (2023). Educational Data Science and Machine Learning: A Case Study on University Student Dropout in Mexico. *Ediciones Universidad de Salamanca*, 1-24.

Naciones Unidas, N. (2023). *Informe de los Objetivos de Desarrollo Sostenible 2023: Edición especial*. Naciones Unidas.

Poulova, P., Klimova, B., & Mikulecká, J. (2019). Data Science—A Future Educational Potential. *International Conference on Multimedia and Ubiquitous Engineering*. Singapore.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining Knowl Discov*, 1-21.

Salas Rueda, R., Eslava Cervantes, A., & Prieto Larios, E. (2020). Teachers' Perceptions about the Impact of Moodle in the Educational Field

Considering Data Science. *Online Journal of Communication and Media Technologies*, 1-10.

Salas-Rueda, R., Salas-Rueda, E., Salas-Rueda, R., & Salas-Silis, J. (2018). Assessment of students about the use of the Alcula application in the educational process through data science and machine learning. *REVISTA GESTION DE LAS PERSONAS Y TECNOLOGÍA*, 12(34), 53-63.

Xiaomei, B., Fuli, Z., Jinzhou, L., & Teng, G. (2021). Educational Big Data: Predictions, Applications and Challenges. *Big Data Research*, 26(100270), 1-12.