



Big data en la educación

Big data in education

Grandes dados em educação

ARTÍCULO DE INVESTIGACIÓN

Hugo Heredia-Mayorga

hugo.heredia@cordillera.edu.ec

hugoheredia79@gmail.com

ORCID: 0000-0001-5534-8934

Instituto Superior Universitario Cordillera, Ecuador

Recibido 18 de noviembre 2020 | Arbitrado y aceptado 12 de diciembre 2020 | Publicado en diciembre 2020

RESUMEN

El estudio buscó determinar la importancia y la aplicabilidad del concepto Big Data como elemento de toma de decisiones en los procesos académicos, administrativos y financieros, el objetivo es describir el papel del Big Data en la educación y conocer el número de investigaciones existente. La metodología que se usó para realizar la revisión sistemática de literatura fue la de Petersen, la cual se clasifica en cuatro etapas. Aplicados los criterios de búsqueda iniciales en los dieron como resultado que existe 69 publicaciones con una baja adopción por parte de las instituciones educativas del Big data, 15 no utilizan herramientas de recolección de datos de forma eficiente lo que dificulta mirar las contribuciones en educación, y 6 publicaciones muestran importantes soluciones planteadas del uso de Big Data. El Big Data es un fenómeno relativamente nuevo en la educación y sus aplicaciones en las actividades académicas permitirán dar forma a un sistema educativo moderno y dinámico.

Palabras clave: Big data; educación; data analytics; data science, educational data analytics

ABSTRACT

The study sought to determine the importance and applicability of the Big Data concept as a decision-making element in academic, administrative and financial processes, the objective is to describe the role of Big Data in education and to know the number of existing investigations. The methodology used to carry out the systematic literature review was Petersen's, which is classified into four stages. Applying the initial search criteria, the result was that there are 69 publications with low adoption by educational institutions of Big data, 15 do not use data collection tools efficiently, which makes it difficult to look at the contributions in education, and 6 publications show important solutions for the use of Big Data. Big Data is a relatively new phenomenon in education and its applications in academic activities will allow to shape a modern and dynamic educational system.

Key words: Big data; education; data analytics; data science, educational data analytics

RESUMO

O estudo buscou determinar a importância e aplicabilidade do conceito de Big Data como elemento decisório em processos acadêmicos, administrativos e financeiros, o objetivo é descrever o papel do Big Data na educação e conhecer o número de investigações existentes. A metodologia utilizada para realizar a revisão sistemática da literatura foi a de Petersen, que é classificada em quatro etapas. Aplicando os critérios de busca inicial, o resultado foi que existem 69 publicações com baixa adoção por instituições de ensino de Big Data, 15 não utilizam ferramentas de coleta de dados de forma eficiente, o que dificulta o olhar para as contribuições na educação, e 6 publicações mostram soluções importantes para o uso de Big Data. O Big Data é um fenômeno relativamente novo na educação e suas aplicações nas atividades acadêmicas permitirão conformar um sistema educacional moderno e dinâmico.

Palavras-chave: Big data; Educação; análise de dados; ciência de dados, análise de dados educacionais

INTRODUCCIÓN

El análisis masivo de datos, también llamado Big Data en la Educación, no fuese posible sin el uso de internet y de herramientas que permitan recolectar la información; de hecho, el Big Data se ha aplicado en muchas industrias, pero carece de elementos sustanciales en la educación lo que ha generado que la revisión de la literatura sea un verdadero desafío.

El Big Data permite priorizar la toma de decisiones sobre la educación, base fundamental del crecimiento y desarrollo dentro de cualquier país; pero, a pesar de aportar significativas ventajas en la industria en la educación aún se muestra el recelo en su aplicación.

Oracle Enterp. Archit ofrece una visión general de la adopción de la Big Data y las capacidades analíticas como parte de una arquitectura de "próxima generación" que

puede satisfacer las necesidades de las instituciones de educación superior (Oracle, 2015); sin embargo, West, (2012) sostiene que las escuelas enfrentan una situación en la que necesitan mejorar la rendición de cuentas global de sus operaciones. En un ambiente de escrutinio público, mediático y político con recursos escasos, las instituciones educativas deben mejorar en la recopilación de datos, el mantenimiento de registros, el análisis y la presentación de informes.

Por su lado, Marconi *et al.* (2013) manifiesta que la Big Data puede realmente mejorar la educación, dando forma a un sistema de educación moderno y dinámico, para que cada estudiante pueda tener el máximo beneficio de eso de manera individual. El sistema educativo se enriquecerá con nuevas formas de aprendizaje, más eficiente y orientado hacia las necesidades de los estudiantes y docentes.

Además, Barbosa (2015) describe que la aplicación de la Big Data en el campo de la educación requiere una investigación mucho más profunda que pueda proporcionar un amplio alcance de comprensión sobre el funcionamiento y habilidades de las instituciones de educación superior en sus cuatro misiones primarias: investigación y análisis de datos, docencia, capacitación y administración; a través de esto elementos a los docentes pueden personalizar los diferentes elementos de cada una de sus cátedras para mejorar el rendimiento de sus estudiantes (Bienkowski *et al.*, 2014).

Entonces, es importante para que las instituciones de educación superior tengan la confianza en el Big Data y proporcionen acceso a los datos mostrar las diversas ventajas que pueden obtenerse de su aplicación. Sin embargo, la aplicación de datos de fuentes diferentes podría ser desafiante,

incluso cuando los datos sean de acceso público, asunto que en las instituciones de educación superior no lo es.

A través de la herramienta del mapeo sistemático abordó una revisión de las investigaciones realizadas por varios autores sobre la pregunta de investigación relacionada a los elementos significativos para mejorar el proceso de aprendizaje de los estudiantes mediante la toma de decisiones. Petersen *et al.* (2008), establece criterios de búsqueda que puedan ser colocados en los diversos repositorios de artículos como Scopus, ACM Digital Library, entre otras.

El objetivo principal de este estudio fue conocer el número de investigaciones existentes sobre la Big Data en la educación, para ello se dividió en cuatro preguntas concretas para obtener resultados detallados, que brinden una visión más general sobre la temática trabajada. Las preguntas que se respondieron en el estudio fueron ¿cuántas publicaciones de la Big Data relacionadas con la educación existen?, si ¿existe evidencia de la adopción de la Big Data en la educación?, ¿qué tipo de contribuciones ha realizado la Big Data en la educación? y ¿qué soluciones del Big Data se han planteado para la educación?

Finalmente, se analizaron los beneficios de aprendizaje de la Big Data dando una breve descripción de cómo estas tecnologías pueden contribuir al sistema educativo, contribuyendo a los profesores y estudiantes a tomar decisiones más específicas sobre los procesos de aprendizaje.

MÉTODO

La metodología que se usó para realizar la revisión sistemática de literatura fue la de Petersen la cual se clasifica en cuatro etapas. A. realizar la búsqueda literaria; B. seleccionar Estudios; C. clasificar los artículos, y D. extraer y realizar la agregación de los datos, con las

variables de los estudios actuales sobre la Big Data aplicada a la educación y sobre los retos y las oportunidades se definió un método para construir clasificaciones y conducir análisis temáticos a efectos de obtener un mapa visual de los conocimientos existentes (Petersen *et al.*, 2008).

Se analizaron los resultados de documentos mediante una categorización que determina Petersen *et al.* (2008) para cada una de las búsquedas y la frecuencia con la que se han realizado publicaciones (Kitchenham *et al.*, 2011). La aplicabilidad de esta metodología de trabajo representa una ventaja con respecto a las revisiones de la literatura tradicional, los estudios comparativos, entre otros, pues permite reducir brechas y unir esfuerzos a través de la revisión de la literatura y realizar una búsqueda sobre este nuevo y relevante tema, sobre su avance y su difusión.

La revisión sistemática de literatura se adapta correctamente al estudio debido a como están planteadas las preguntas para conocer el estado de las investigaciones sobre las oportunidades y los desafíos del Big Data en la educación superior. Esto es un punto de partida para la realizar futuras exploraciones y resultados sobre cada una de las respuestas encontradas a los planteamientos realizados.

Petersen *et al.* (2008) sugiere cinco etapas para la realización de un proceso de mapeo sistemático de revisión, de los cuales se tomó cuatro que se detallan:

1. Fuentes de datos y estrategias de búsquedas

Para el estudio se planteó una estrategia de búsqueda considerando los aspectos inherentes a cada uno de los términos relacionados con la problemática a trabajar, es decir, sobre la Big Data en la educación. Los distintos procesos de revisión sistemática generalmente causan miedo e impresión por el desconocimiento que

se tiene sobre su aplicabilidad y uso, generando que sus resultados sean confusos e imprecisos. Aun no se encuentra totalmente elementos que permitan evidenciar claramente la utilidad que la Big Data puede dar a la educación, sobre todo en lo que respecta a la educación superior.

Por este motivo las cadenas de búsqueda se determinan de manera sistemática buscando elementos relacionados con el tema de la investigación (Aurum *et al.*, 2002), permitiendo de cierta manera depurar el vocabulario y conocer el uso de sinónimos.

Los motores de búsqueda seleccionados para realizar la búsqueda de la documentación son las bases de datos Scopus, ACM Digital Library, Springer, IEEE Digital Library, Wiley Online Library, World Scientific Net, mismas que ofrecen una amplia cobertura de publicaciones indexadas.

Para las búsquedas se contemplan las publicaciones relacionadas con las siguientes áreas de conocimiento: Big Data, Analítica de Datos, Data Science, Machine Learning, Educational Dataminig, Educational Data Science, E-learning Anatytic, cubriendo los

trabajos presentados desde el año 2013 hasta el año 2016 donde se llevó a cabo la investigación.

Las cadenas de búsquedas utilizadas para el análisis del trabajo han sido definidas de acuerdo con las sugerencias de Beecham *et al.* (2006), donde cada una de ellas está basada a partir de la pregunta de investigación, es decir en el formato Población, Intervención, Comparación, Salida y Contexto (PICOC por sus siglas en inglés), misma que aplica la identificación de palabras relacionadas y sinónimos de los términos de la pregunta de investigación. Para la vinculación de las palabras se debe utilizar las palabras OR, para vincular la población, intervención, contexto se debe trabajar con el operador AND.

Se aplicó los criterios de búsqueda que se realizarán, lo primero que se hizo fue definir los criterios iniciales de búsqueda determinando los sinónimos de los términos identificados en la pregunta de investigación, tal como lo demuestra la Tabla 1, dichos términos se encuentran en el idioma inglés pues lo motores de búsqueda seleccionados así lo contemplan.

Tabla 1. Criterios de búsqueda iniciales

PICOC	Término	Sinónimo/Término relacionado
Población	Education	Scholarship, Schooling, Educational, Learning, Higher education,
Intervención	Challenge and opportunities	Challenge, opportunities, advantage, theat, desventaje, problems, benefit.
Contexto	Big Data	Data Analytics, Data Science, Educational Data Science, Learning Analytics, Educational Dataminig, Dataminig, Big Data in Education

2. Selección de estudios

Para la selección de los estudios que se refieran a cada una de las preguntas del estudio se definió dos criterios: la inclusión y la exclusión propias del mapeo sistemático.

3. Criterios de inclusión

- CI1: Se debe tener en cuenta todos aquellos trabajos que tratan sobre la Big Data enfocados en la educación, Analítica de Datos, Data Science, Machine Learning, Educational Datamining, Educational Data Science, E-learning Analytic.
- CI2: Se debe considerar aquellos artículos que tengan relación entre el título, objetivos, resumen y palabras claves entre sí.
- CI3: Los artículos deberán identificar las palabras claves trabajadas en el contexto.

4. Criterios de exclusión

- CE1: Se decide quitar aquellos trabajos que no traten de la Big Data enfocados en

la educación, Data Analytics, Data Science, Machine Learning, Educational Datamining, Educational Data Science, E_learning Analytic.

- CE2: Los resúmenes de los artículos no cumplen con los objetivos de la pregunta de investigación.
- CE3: Se descartan artículos que no tengan nada que ver con educación, o sinónimos relacionados.

Como la información con la que se empezó a trabajar es desconocida y de gran volumen, se realizó una búsqueda exploratoria para obtener una primera revisión sobre la temática a realizar.

RESULTADOS

Los primeros resultados obtenidos aplicando los filtros establecidos en la Tabla 2, y en función de los criterios de inclusión y exclusión; en las búsquedas en cada uno de los motores de búsqueda se identificaron los siguientes resultados:

Tabla 2. Resultados de la búsqueda en los repositorios digitales

Base de Datos Digital	Numero de publicaciones
ACM Digital Library	132
Scopus	218
Springer	110
IEEE Digital Library	86
Wiley Online Library,	40
World Scientific Net	14

En función de lo que dice Van Rijsbergen y Jardine (1971) para el análisis de la información en la primera búsqueda determina high recall (HR), para dar a conocer toda la información relevante e información de poca utilidad para la

investigación. El segundo, low precisión (LP), ocurre cuando la cantidad de la información recuperada es totalmente relevante en la búsqueda y en los resultados obtenidos, tal como se visualiza en la Tabla 3.

Tabla 3. Análisis de acuerdo con parámetros de Rijsbergen

Base de Datos Digital	Numero de publicaciones	Motivo	Aplican	Precisión
ACM Digital Library	132		40	30%
Scopus	218		15	7%
Springer	110		10	9%
IEEE Digital Library	86		22	26%
Wiley Online Library	40	HR, LP		
World Scientific Net	14	LP		

Los datos que se observan en la Tabla 3, permiten observar la precisión en porcentaje de los resultados comprendidos en los criterios de inclusión y exclusión, el resultado se determina a partir de la unión de la información sin duplicados correspondientes a las cadenas determinadas en la Tabla 1.

A. Clasificación de Artículos

Para realizar la clasificación de los artículos que fueron resultado de haber ejecutado los criterios de búsquedas de la Tabla 1 en cada uno de los motores de búsqueda, se definió los criterios para establecer la pertenencia a una categoría adaptada a la temática trabajada, los criterios bajo los cuales se realizó la clasificación son:

- Tipo de Investigación: este tipo de clasificación es adaptado del trabajo presentado por Rolland *et al.* (2006), que se utiliza para la clasificación de búsqueda en ingeniería de requerimientos.
- Área en Educación: clasifica los diferentes resultados encontrados en el ámbito de educación y todo lo que esté relacionado con ello.
- Tipo de contribución: para este tipo de clasificación se debe establecer el tipo de contribuciones que realiza cada una de las publicaciones encontradas, para ello se basa en el trabajo realizado por Petersen *et al.*

(2008), el resumen de cada publicación es resumido para destacar las principales palabras clave que describieron la contribución de la investigación, mismas que fueron añadidas y los sinónimos se fusionaron para simplificar los resultados y promover una mejor visibilidad de las tendencias.

- Tipo de Análisis: para realizar la clasificación por el tipo de análisis se definirá un esquema como lo que propone Delen y Demirkan (2013), que es una taxonomía de alto nivel para realizar análisis de negocios.

B. Extracción y análisis de datos

Las guías de Petersen *et al.* (2008) sugieren la exploración de algunas partes del artículo sólo en aquellos casos donde el abstract no se encuentra bien estructurado o es impreciso.

Para esta revisión sistemática se decidió realizar la lectura completa de cada artículo para poder responder todas las preguntas planteadas, ya que el solo hecho de analizar títulos y abstracts no basta.

También se incluyen en la extracción y síntesis aquellos trabajos que tratan el tema de la adopción directamente.

Se determinó una plantilla para la extracción de los datos resultantes de cada cadena de búsquedas consideradas en el trabajo en donde se registran los artículos encontrados junto con toda la información bibliográfica

correspondiente, agregando el criterio de clasificación anteriormente expuestos.

Posteriormente se establece una plantilla única quitando los artículos duplicados y establecer las frecuencias y análisis correspondiente.

Luego de haber aplicado cada uno de los criterios de inclusión y exclusión a las distintas cadenas de búsqueda, la se obtiene un total de 15 artículos que aplican al tema de

investigación, el resultado presentado en el mapeo sistemático de la literatura se presenta a continuación como respuestas a las interrogantes planteadas en el trabajo.

La respuesta a la pregunta RQ1 sobre las publicaciones de Big data relacionadas con la Educación se presentan en función de los resultados obtenidos en las búsquedas aplicando cada uno de los criterios y cadenas de búsqueda.



Figura 1. Crecimiento de la Publicaciones SCOPUS

La Figura 1 ilustra el crecimiento de las publicaciones relacionadas Big Data, Data Analytics, Data Science, Machine Learning, Educational Dataminig, Educational Data Science, E_learning Anatyctic en la educación. En el motor de búsqueda Scopus. Las publicaciones

analizadas es desde el año 2013 hasta el año 2016 de 69 publicaciones evidenciando un crecimiento de 110% al 2014, pero para los años 2015 se observa un decrecimiento considerable de 110% del 2104 al 41 % al 2015 y al 2016 apenas alcanza al 7%.

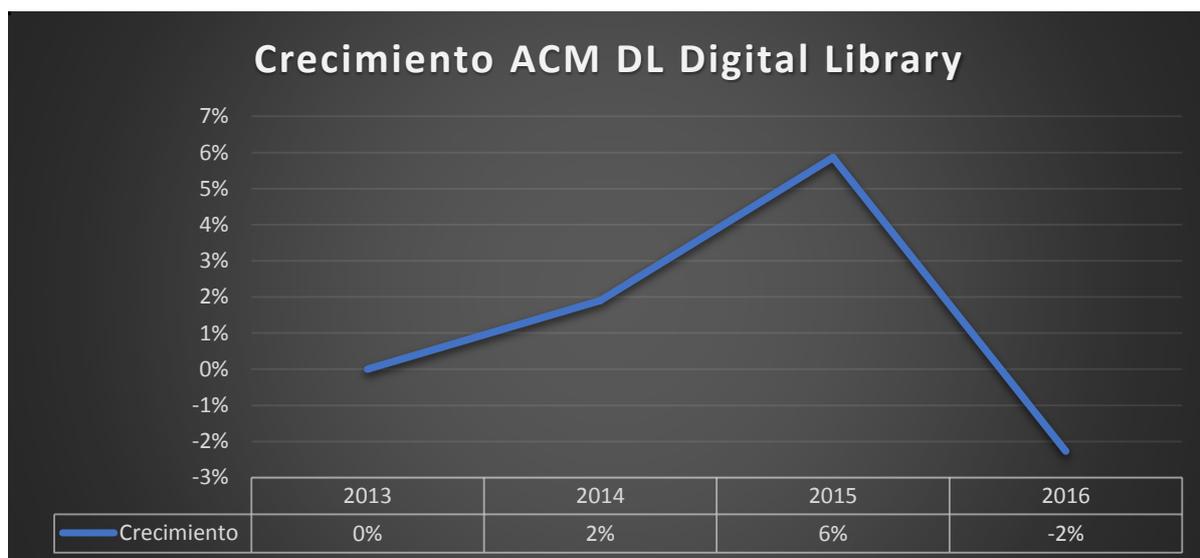


Figura 2. Crecimiento de la Publicaciones ACM DL Digital Library

La Figura 2 ilustra el crecimiento de las publicaciones relacionadas Big Data, Data Analytics, Data Science, Machine Learning, Educational Dataminig, Educational Data Science, E_learning Anatytic en la educación. En el motor de búsqueda ACM DL Digital Library. Las publicaciones analizadas es desde el año 2013 de 9063 publicaciones evidenciando un crecimiento de 2% al 2014, para el año 2015 se observa un crecimiento considerable de 2% del 2104 al 6 % al 2015 mientras que en el 2016 apenas alcanza 132 publicaciones.

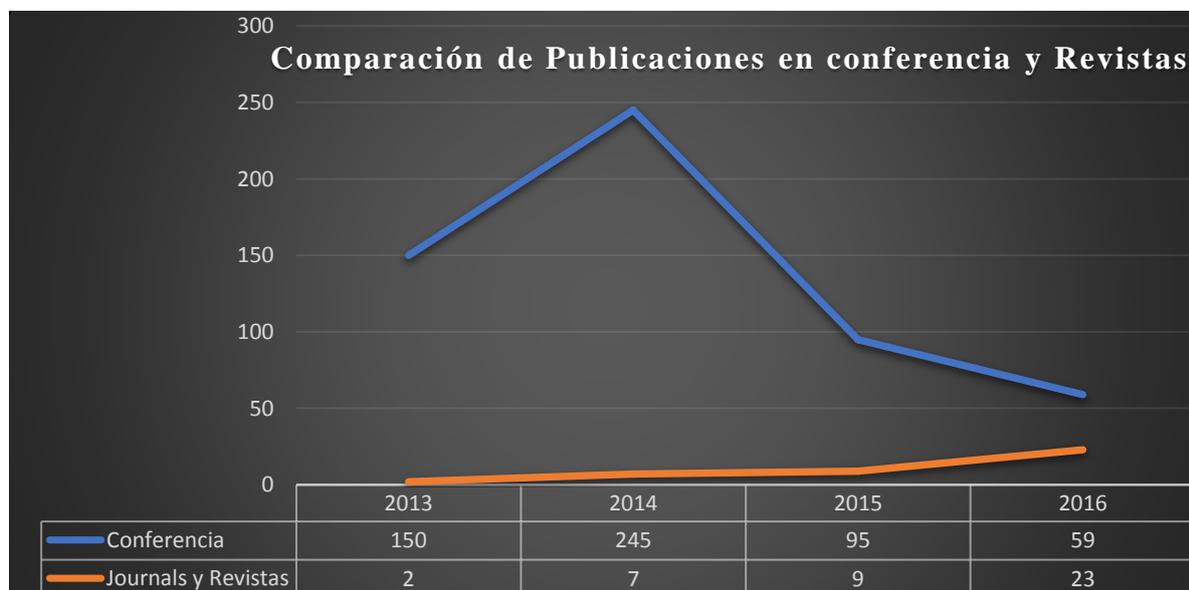


Figura 3. Comparación Publicaciones Conferencias Vs Revistas

En la Figura 3 se presenta un desglose de publicaciones por revistas y conferencias, las publicaciones de la conferencia fueron mayores que las de las revistas para cada año que se realizó las búsquedas, es importante observar que a partir del año 2015 las publicaciones en las revistas comienzan a subir, pero aún no es tan considerable, en cambio las publicaciones en las conferencias empiezan a disminuir a partir del año 2015. Si analizamos en términos de correlación a simple vista se observa que no guardan en ningún momento una correlación directa en las publicaciones.

La distribución de las publicaciones de conferencias y las revistas; la mayor cantidad es en la revista de la IEEE, en lo referente a la Big Data, Data Analytics, Data Science, Machine Learning, Educational Datamining, Educational Data Science, E_learning Analytic en la educación, considerándose las más actuales a las publicaciones realizadas en el año 2014.

Con respecto a la pregunta RQ2: sobre si existe evidencia de la adopción de la Big Data en la educación se confirma que existe una baja adopción en el análisis de los datos en la información por parte de las instituciones educativas.

Para que las instituciones de educación adopten las técnicas de la Big Data y que los padres y maestros puedan evaluar lo que está sucediendo en el aula, deben realizar la formulación de políticas que permitan evaluar el aprendizaje y el logro en los procesos académicos.

También hay autores como Gutiérrez *et al.* (2014) que manifiestan textualmente en su trabajo que, en los próximos años, el análisis de la Big Data en entornos educativos irá

progresivamente alcanzando su etapa de madurez proporcionando nuevos campos de estudio y análisis.

En lo que respecta a la pregunta RQ3 se analiza las contribuciones que ha realizado la Big Data en la educación, West (2012) sostiene en su trabajo que las escuelas enfrentan una situación en las que necesitan mejorar la rendición de cuentas global de sus operaciones. Las instituciones educativas deben mejorar en la recopilación de datos, el mantenimiento de registros, el análisis y la presentación de informes.

Piety *et al.* (2014) en su artículo manifiesta que para trabajar con la Big Data y la Data Analytics para medir el crecimiento asociado con datos educativos en un campo emergente, se considera como socio técnico y transdisciplinario que para trabajar requerirá de una combinación de habilidades técnicas y sociales; Una aptitud para la ingeniería y también la comprensión profunda del complejo mundo, de la práctica educativa y el aprendizaje.

También existen afirmaciones que indican que la Big Data puede proporcionar más oportunidades, nuevas experiencias de aprendizaje para niños, jóvenes y adultos. Por lo tanto, los estudiantes pueden compartir información con instituciones que pueden ampliar sus conocimientos y habilidades. Además, los institutos, las universidades pueden ayudar a preparar su futuro Marconi *et al.* (2013). Por último, existen factores de costo a considerar.

Para encontrar respuestas a la pregunta RQ4 en lo referente a las soluciones que la Big Data ha planteado para la Educación se resume las más importantes en la Tabla 4.

Tabla 4. Soluciones planteadas resultantes en el mapeo sistemático

Solución Planteada	Referencia
<p>En la mejora de la gestión educativa, al desarrollo de nuevos métodos para la enseñanza y el aprendizaje, la creación de nuevas carreras y opciones profesionales, así como en la generación y almacenamiento de acervos digitales que constituyen el producto de años de actividad académica, docente y de investigación.</p>	<p>(Salazar Argonza, 2016)</p>
<p>La analítica predictiva puede identificar qué estudiantes no pueden terminar su grado a tiempo o incluso entregar en asignaciones individuales, ya se está viendo en el Reino Unido a través de la herramienta OU Analyse tool</p>	<p>(Shacklock, 2016),(Argotty y Herenádez, 2013), (Tabares y Hernández, 2014)</p>
<p>Activate Learning es mejorar la calidad del rendimiento estudiantil, la retención y entender por qué los estudiantes no logran tener éxito para que esta información pueda ser proporcionada para que se puedan hacer mejoras.</p>	<p>(Sclater <i>et al.</i>, 2014), (Ducange <i>et al.</i>, 2017)</p>

Discusión

Del resultado del mapeo sistemático realizado se obtuvieron observaciones y contribuciones de la Big Data, Data Analytics, Data Science, Machine Learning, Educational Dataminig, Educational Data Science, E_learning Analytic realizadas por otros autores, lo que concuerda con los resultados del estudio

En primer lugar, el análisis de la Big Data en entornos educativos crecerá de tal manera que permitirá a las Ciencias de la Computación articular elementos que permitan mayor capacidad de almacenamiento y procesamiento de datos, a diferencia de lo que ocurría con el análisis mediante encuestas, entrevistas, transformándose los datos e información un recurso escaso a un recurso abundante, o incluso sobre abundante. Desde un punto de vista metodológico, las técnicas de predicción e inferencia estadística serán cada vez más utilizadas, abordando no sólo cuestiones técnicas, sino también implicaciones sociales, ingeniería social y entornos en donde interactúa el estudiante en la actualidad.

Sin embargo, el camino de esta nueva era apenas ha comenzado y hay muchas dificultades como la falta de personal experimentado en la ciencia de la Big Data y Data Analytics, los docentes y estudiantes deben aceptar y utilizar estas nuevas herramientas (Podesta *et al.*, 2014).

También se mencionan contribuciones en forma de mejoras, lo que concuerda con las afirmaciones y evidencias encontradas en los artículos sobre el problema planteado, sin embargo, existen acercamientos con nuevas tecnologías y propuestas, pero nunca llegarán a salir del ámbito industrial y las que si lo hacen generan muy poca información empírica sobre su aplicabilidad en otras áreas.

De esta forma la Big Data será utilizada en una variedad de sectores industriales, pero la falta de formación y la existencia de guías y evidencias que contribuyan a la toma de decisiones acertadas por los administradores y autoridades de los centros de educación es una de las limitantes de su utilización en el ámbito educativo.

El procesamiento de los resultados del mapeo sistemático de la literatura confirma que el tema no toma aun el interés necesario para la comunidad educativa y que se debe trabajar más en él. Otro resultado de este análisis es que las instituciones de educación albergan la inseguridad sobre los pocos resultados evidentes de la adopción de la Biga Data en la educación y la forma en la que su información este siendo analizada.

Otro de los aspectos asociados a la falta de elementos que puedan articular la Big Data a la educación está asociado con la formación de profesionales y capacitación en general, la falta de entrenamiento, las curvas de aprendizaje de los estudiantes, no refleja resultados convincentes que permitan y coadyuven en forma definitiva a tomar la decisión de adoptar nuevos paradigmas para mejorar la educación.

Dichas soluciones, así como también cualquier otra que surja de los análisis hallados, deberán ser estudiadas en mayor profundidad. La principal dificultad de la incorporación de algunas de estas sugerencias es que se desconoce qué implicaciones pueden llegar a tener en una cultura arraigada y mal concebida de lo que llamamos educación, de pasar de una educación tradicionalista, de tener docentes que dictan clases a docentes que se conviertan en verdaderos guías y personalicen el aprendizaje de sus estudiantes para que de esta manera se pueda alcanzar los niveles académicos deseados, y los profesionales que se buscan para alcanzar el desarrollo de los países.

A partir del año 2017 en adelante ha existido un incremento significativo sobre el análisis del Big Data y Data Analytics en la educación, estudio como el de Garzón-Benalcazar *et al.* (2018) en a que hablan el impacto u beneficios de la aplicación del Big Data y Data Analytics en la educación cuyas conclusiones permiten tener mayores criterios sobre la aplicabilidad de los

datos debido a las mayores facilidades tecnológicos que se presentan día a día.

CONCLUSIÓN

Como se puede ver el Big Data puede realmente mejorar la educación. Puede dar forma a un sistema educativo moderno y dinámico, que cada estudiante individual pueda tener el beneficio máximo de eso.

Además, los profesores tienen herramientas valiosas, que no lo tenían antes, lo que puede hacer que sus decisiones sean más específicas y puedan elegir una gran variedad de nuevos métodos de aprendizaje.

Por lo tanto, la Big Data y el Data Analytics están realmente diseñados para innovar los procesos, la toma de decisiones en la industria moderna, incluida la educación. Esta nueva era de los Datos trae consigo dificultades en la aplicación de métodos de análisis predictivo, que enriquecen la nueva forma de aprendizaje, falta de personal experimentado en el análisis de datos masivos, transformando la educación de algo masivo a algo totalmente personalizado, mejorando la selección de los materiales de trabajo, potencializar una mejor toma de decisiones sobre las estrategias didácticas y metodológicas que se deben aplicar, apoyados con una orientación vocacional enfoca a mejorar la elección de la carrera por parte d los estudiantes.

REFERENCIAS

- Argotty, I., y Herenádez, G. (2013). Construcción de Modelos Matemáticos para determinar el nivel de deserción en los programas de pregrado de la Universidad Mariana. In *UNIVERSIDAD MARIANA*
- Aurum, A., Petersson, H., y Wohlin, C. (2002). State-of-the-art: software inspections after 25 years. *Software Testing, Verification and Reliability*, 12(3), 133-154

- Barbosa, J. G. (2015). *La verdadera transformación educativa pasa por big data*. La Verdadera Transformación Educativa Pasa Por Big Data. <http://aunclidelastic.blogthinkbig.com/la-verdadera-transformacion-educativa-pasa-por-big-data/>
- Beecham, S., Baddoo, N., y Hall, T. (2006). Protocol for a Systematic Literature Review in Software Engineering. *University of Hertfordshire*
- Bienkowski, M., Feng, M., y Means, B. (2014). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. In *Educational Improvement Through Data Mining and Analytics* (pp. 1–60). Nova Science Publishers, Inc
- Delen, D., y Demirkan, H. (2013). Data, information and analytics as services. *Decis Support Syst* 55, 359–363
- Ducange, P., Pecori, R., Sarti, L., y Vecchio, M. (2017). Educational big data mining: How to enhance virtual learning environments. In *Advances in Intelligent Systems and Computing* (Vol. 527). https://doi.org/10.1007/978-3-319-47364-2_66
- Garzón-Benalcazar, J., Tejada-Escobar, F., Murieta-Marcillo, R., y Villao-Santos, F. (2018). Big Data en la Educación: Beneficios e impactos de la Analítica de Datos. *Revista Científica y Tecnológica UPSE*, 5(2), 88–96
- Gutiérrez, D., Reina, D. G., Toral, S. L., y Barrero, F. (2014). plataformas educativas Metodologías de Análisis de los Big Data en las Plataformas Educativas. *ReseachGate, November*
- Kitchenham, B. A., Budgen, D., y Pearl Brereton, O. (2011). Using mapping studies as the basis for further research—A participant-observer case study. *Information and Software Technology*, 53(6), 638–651. <https://doi.org/10.1016/j.infsof.2010.12.011>
- Marconi, K., Dobra, M., y Thompson, C. (2013). The Use of Big Data in Healthcare. *Big Data and Business Analytics*, 11(5), 229–248. <https://doi.org/doi:10.1201/b14700-15>
- Oracle. (2015). Improving Higher Education Performance with Big Data. In *Oracle Enterprise Architect Ure White Paper* (Issue April). <http://www.oracle.com/us/technologies/big-data/big-data-education-2511586.pdf>
- Petersen, K., Feldt, R., y Mujtaba, S. (2008). Systematic mapping studies in software engineering. *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, 1–10
- Piety, P. J., Hickey, D. T., y Bishop, M. J. (2014). Educational data sciences - Framing emergent practices for analytics of learning, organizations, and systems. *ACM International Conference Proceeding Series*, 193–202. <https://doi.org/10.1145/2567574.2567582>
- Podesta, J., Pritzker, P., Moniz, E. J., Holdren, J., y Zients, J. (2014). Big Data: Seizing Opportunities. *Executive Office of the President of USA, May*, 1–79. <https://doi.org/10.5121/ijgca.2012.3203>
- Rolland, C., Wieringa, R., Maiden, N., y Mead, N. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requir Eng*, 11, 111–112
- Salazar Argonza, J. (2016). Big data. *Revista Digital Universitaria*, 17(1), 1–16. <http://www.revista.unam.mx/vol.17/numero1/art06/>
- Sclater, N., Henry, J., y Building, B. (2014). *Learning analytics The current state of play in UK higher and further education*
- Shacklock, X. (2016). *From Bricks to clicks. The Potential of Data and Analytics in Higher Education*. http://www.policyconnect.org.uk/hecsites/site_hec/files/report/419/fieldreportdownload/frombrickstoclicks-hecreportforweb.pdf
- Tabares, L. F., y Hernández, J. F. (2014). Big Data Analytics : Oportunidades , Retos y Tendencias. *Universidad de San Buenaventura*, 20

Van Rijsbergen, C. J., y Jardine, N. (1971). The use of hierarchic clustering in information retrieval. *Information Storage and Retrieval*, 7(5), 217-240. [https://doi.org/10.1016/0020-0271\(71\)90051-9](https://doi.org/10.1016/0020-0271(71)90051-9)

West, D. (2012). Big Data for Education: Data Mining, Data Analytics, and Web

Dashboards. *Governance Studies at Brookings*, September, 11. http://www.brookings.edu/~media/research/files/papers/2012/9/04_education_technology_west/04_education_technology_west