

16

**ANALÍTICA ACADÉMICA: NUEVAS HERRAMIENTAS
APLICADAS A LA EDUCACIÓN.**

**ACADEMIC ANALYTICS: NEW TOOLS APPLIED TO
EDUCATION.**

Leonardo Emiro Contreras Bravo ¹
José Ignacio Rodríguez Molano ²
Héctor Javier Fuentes López ³

Universidad Distrital Francisco José de Caldas (Colombia)

¹ Ingeniero, Estudiante de doctorado en ingeniería, Docente de planta. Facultad de ingeniería. Universidad Distrital Francisco José de Caldas (Colombia). lecontrerasb@udistrital.edu.co. ORCID: <https://orcid.org/0000-0003-4625-8835>

² Ingeniero, Doctor en Ingeniería. Docente de planta. Facultad de ingeniería. Universidad Distrital Francisco José de Caldas (Colombia). jirodriguez@udistrital.edu.co ORCID: <https://orcid.org/0000-0003-2581-277X>

³ Economista, Estudiante de doctorado en Economía. Docente de planta. Facultad de ingeniería. Universidad Distrital Francisco José de Caldas (Colombia). hjfuentesl@udistrital.edu.co ORCID: <https://orcid.org/0000-0001-6899-4564>

RESUMEN

La analítica de datos es un campo nuevo que ha permeado la educación superior mediante la incursión de herramientas matemáticas, la estadística, la minería de datos y el aprendizaje automático. Inicialmente se presenta una fundamentación teórica relacionada con la analítica aplicada a la educación, analítica académica y sus enfoques. Posteriormente se plantea una metodología cuyo propósito es la revisión referencial de los últimos cinco años referente al campo de la analítica en educación y especialmente en lo que concierne a la analítica académica, con el fin de identificar aspectos relativos al crecimiento de este enfoque y sus campos de aplicación, enfocados a la educación superior. Los resultados muestran que los investigadores se han enfocado en los últimos años a trabajar en el desarrollo de modelos que permitan comprender aspectos de la vida académica del estudiante, docentes e instituciones (rendimiento académico, tasa de deserción y tasa de graduación en su respectivo orden) que permitan la elaboración y toma de decisiones acertadas.

PALABRAS CLAVE: Analítica; analítica académica; aprendizaje automático; educación en ingeniería

SUMMARY

Data analytics is a new field that has permeated higher education through the foray into mathematical tools, statistics, data mining, and machine learning. Initially, a theoretical foundation related to applied analytics in education, academic analytics and their approaches is presented. Subsequently, a methodology is proposed whose purpose is the referential review of the last five years regarding the field of analytics in education and especially in what concerns academic analytics, in order to identify aspects related to the growth of this approach and its fields application, focused on higher education. The results show that researchers have been concerned in recent years to work on the development of models that allow understanding aspects of the academic life of the student, teachers and institutions (academic performance, dropout rate and graduation rate in their respective order) that allow the development and making of correct decisions.

KEYWORDS: Analytics; academic analytics; machine learning; engineering education

INTRODUCCIÓN

El sector de las tecnologías de la información y comunicación (TIC) se muestra actualmente como líder en la manipulación de datos provenientes de diferentes medios como redes sociales, plataformas tecnológicas y dispositivos electrónicos de comunicación que requieren emplear la información capturada o adquirida para ser transformada y suministrar informes estadísticos descriptivos o tal vez informes sobre modelos predictivos aplicables a diversas áreas del saber. Tal como lo plantean (Greller & Drachsler, 2012) quienes plantean las ventajas que la introducción de las TIC ha causado en el sector educativo, de hecho, su argumento se basa en la importancia de la tecnología para desarrollar procesos que antes no se podían realizar, como es el caso de la introducción de analíticas para determinar múltiples factores tales como el monitoreo del desempeño de los estudiantes, información actualizada y procesada acerca del progreso del alumno, entre otros.

En el campo de la educación se han estado generando una cantidad creciente de información (datos) producto de la iteraciones de los diferentes actores del proceso (docente, institución y estudiantes) por medio del desarrollo de herramientas informáticas y bases de datos (Schroeder, 2009). Datos que son almacenados y poco son usados con el fin de mejorar los factores

importantes de la eficiencia del proceso educativo tales como el rendimiento académico del estudiante (indicador de éxito o fracaso del estudiante) (Estrada & Quintero, 2015), parámetros de retención (diferencia entre número de estudiantes que ingresan en primer semestre y los graduados por año) (Salcedo, 2010) y la orientación académica (factores que afectan la posibilidad de abandono de un programa de educación) (L. Contreras & Tarazona, 2018).

La analítica está siendo usada por las universidades para monitorear el progreso de los objetivos institucionales. Existen diversos métodos utilizados en la predicción del rendimiento académico, algunos de estos métodos utilizan datos de plataformas en línea de aprendizaje, incorporadas por las universidades (Wang et al., 2018). Otros toman los datos históricos en la utilización de algunos métodos de estudio ya que muchas veces estos no se tienen, es por ello que estos se deben recolectar, depurar y analizar, en los cuales toma mucha relevancia el tamaño de la muestra (Lasarte et al., 2020).

En el presente documento se estudiarán aspectos relacionados con la analítica académica, es por ello que es importante diferenciarla de la analítica del aprendizaje, la principal diferencia que presentan estos dos campos, es que la primera de estas (analítica académica) se centra en el apoyo para los procesos relacionados con la institución educativa a través del estudio del rendimiento académico de los estudiantes, niveles de deserción estudiantil, entre otras variables; mientras que la analítica del aprendizaje se centra en la comprensión, estudio y predicción del aprendizaje de los estudiantes apoyados en el estudio virtual (Viberg et al., 2018) El propósito de la revisión referencial consiste en examinar la literatura de los últimos cinco años referente al campo de la analítica en educación y especialmente en lo que concierne a la analítica académica, con el fin de identificar aspectos relativos al crecimiento de este enfoque y sus campos de aplicación, enfocados a la educación superior.

FUNDAMENTACIÓN

Para comprender esta temática se describen a continuación los principales conceptos relacionados en esta investigación referencial: analítica y analítica académica.

Analítica en educación

La analítica de datos es un campo nuevo que ha surgido en la educación superior a raíz de la incursión de una serie de herramientas matemáticas y estadísticas inicialmente, que junto con el desarrollo de las ciencias de la computación y de la inteligencia artificial han permitido la toma de decisiones basadas en datos con fines operativos en los diversos niveles de la educación (Baepler & Murdoch, 2010a), aplicándose no solo a los problemas de enseñanza y aprendizaje de los estudiantes sino también a problemas de retención, graduación y deserción. Incursión que puede mostrarse en diferentes etapas según (Murnion & Helfert, 2013) quienes manifiestan que la analítica puede generar valor a partir del análisis realizado (**Figura 1**).

Analítica	Descriptiva (Desempeño del total de las actividades)	Informes, cuadros, tablas, graficas estadísticas <hr/> Pregunta Clave: ¿Que pasó? <hr/> Herramienta: Software pra analisis general, hojasde cálculo
	Diagnóstica (Desempeño pasado para analizar la información)	Tendencias o relaciones causales entre las variables y los resultados <hr/> Pregunta Clave: ¿Porque sucedió? <hr/> Herramienta: Estadística, matematicas, ciencias computacionales
	Predictiva (Anticiparse a los comportamientos basándose en relaciones históricas entre las variables)	Predicciones, segmentaciones, optimizaciones, simulaciones <hr/> Pregunta Clave: ¿Que podría pasar? <hr/> Herramienta: Minería de datos, aprendizaje automático

Figura 1. Panorama de analítica de datos. Fuente: Adaptado de (Brownlee, 2020)

La analítica de datos educativos se puede sintetizar según (Clow, 2013) y (Siemens, 2013) en los siguientes pasos:

- **Pregunta de investigación:** Consiste en definir el problema y en tener claridad de que es lo que se pretende resolver, a quien o quienes beneficiará dicha solución.
- **Recolección de datos:** Los datos pueden proceder de diferentes fuentes (sistema académico, de bibliotecas, de gestión de aprendizaje) y en múltiples formatos (hojas de cálculo, registros en papel, bases de datos).
- **Reportar información:** Plantea el desarrollo de informes y análisis de la información con el fin de identificar tendencias, patrones y correlaciones éntrelas variables analizadas.
- **Análisis y Predicción:** Es una etapa importante de la analítica que involucra el desarrollo de modelos usando herramientas estadísticas, de minería de datos o de aprendizaje automático.
- **Acción/ realimentación:** Conlleva a la toma de decisiones por parte de docentes, estudiantes e instituciones basadas en la evidencia (datos).

Al hablar de analítica hay que tener claridad que este término involucra diversas técnicas, herramientas y métodos que por lo general se traslapan, tales como como son la estadística (Baepler & Murdoch, 2010b), inteligencia de negocios (Argonza, 2016), minería de datos (García-González et al., 2019) y aprendizaje automático (Burman & Som, 2019). Métodos que tiene por tarea desafiante el recopilar, limpiar, transformar y analizar múltiples datos provenientes de diversas fuentes contribuyendo al futuro, con especial referencia al contexto de la educación desarrollando modelos predictivos para abordar problemas críticos en la educación superior (**Figura 2**). El uso de Analítica en educación posee dos grandes enfoques: la Analítica del Aprendizaje que se orienta al proceso de enseñanza y aprendizaje y sus diferentes actores (Ferguson, 2012) ; y la Analítica Académica que se centra en los indicadores que permiten optimizar la labor social de la educación permitiendo comparar resultados entre instituciones (Daniel, 2015b). Involucra análisis en diferentes niveles de la institución superior tales como:

departamentos, proyectos curriculares, facultades y hasta universidades

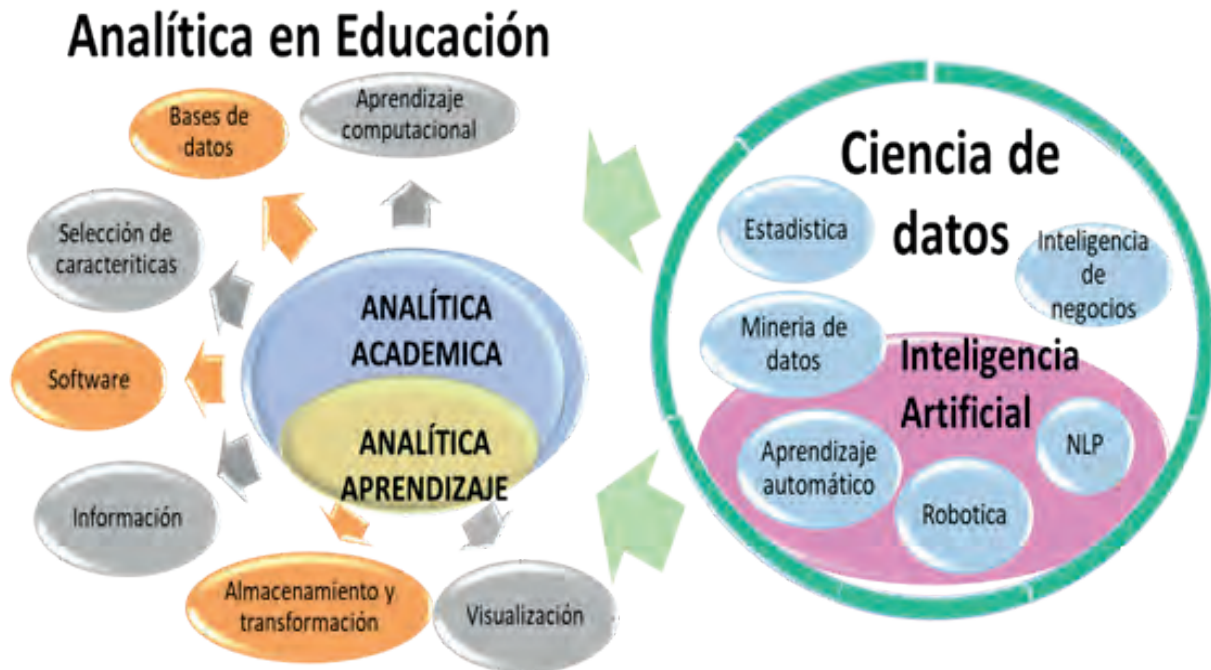


Figura 2. Influencia de herramientas tecnológicas en la analítica de datos educativos. Fuente: Autor

Definición de Analítica Académica.

Diferentes autores han tratado de definir el concepto de analítica académica. (Norris & Lefrere, 2011) manifiestan que la analítica académica consiste en diversos procesos de evaluación y análisis de datos que permiten medir, mejorar y comparar el desempeño de individuos, programas o instituciones. Es un análisis enfocado en aspectos académicos, relativos al éxito de los estudiantes según (Norris & Lefrere, 2011).

La analítica académica parte de la recopilación de la información y su análisis por medio de un enfoque denominado minería de datos que posee diversos métodos de aplicación, entre los cuales algunos de ellos se tratan de la búsqueda de patrones sobre el conjunto de datos que se desea analizar (Daniel, 2015a). Por tanto (Iten et al., 2008) son autores que manifiestan que la minería de datos aplica al campo de la educación en el sentido de aplicar una serie de algoritmos que permiten fortalecer el cumplimiento de los objetivos de las instituciones de educación superior, cuya finalidad es lograr los aspectos como la mejora de rendimiento académico, disminuir la tasa de deserción e incrementar la tasa de graduación (Iten et al., 2008).

La analítica académica tiene por objetivo permitir a las instituciones educativas, profesores, estudiantes y demás entes involucrados en el proceso educativo monitorear y medir el logro de los objetivos institucionales, a través de la identificación de variables relacionadas con los logros de los estudiantes (créditos cursados, promedio académico, porcentaje de asistencia a clases, entre muchas otras), que permita estudiar y predecir el comportamiento de los estudiantes con respecto a otras variables de resultado, tales como: nivel de deserción y riesgo. (García Tinisaray, 2015), (Lauría et al., 2013).

Enfoque de Analítica Académica.

El proceso de la analítica académica, se puede describir en 5 pasos fundamentales: 1) inicialmente se recolecta información, que puede ser de fuentes externos o internas de la institución educativa, por ejemplo el historial académico de los estudiantes, 2) se identifican los datos buscando tendencias 3) se realizan modelos predictivos teniendo en cuenta la información recolectada, 4) dada la información obtenida de los modelos, la instituciones actúan y por último, 5) se perfecciona o mejora el proceso educativo. *(García Tinisaray, 2015)*

En resumen, *(Gómez, 2015)* conceptúa que la medición del rendimiento educativo por medio de la analítica incluye dos aspectos importantes que se complementan. Por un lado, debe considerarse el rendimiento operativo (sistemas administrativos y de apoyo) y por otro el rendimiento académico (diseño y ejecución de estrategias académicas para lograr experiencias de aprendizaje, los resultados y las competencias de la vida real). Pero que así mismo el éxito de la analítica académica se ve influenciada por aspectos claves tales como; armar un equipo líder del proceso comprometido en tomar acciones sobre la base de los resultados encontrados; la conformación de un equipo que posea el conocimiento y manejo de las herramientas necesarias para recopilar, transformar los datos y crear modelos predictivos; y por ultimo poseer la tecnología adecuada para el desarrollo de la analítica, dentro de lo cual se considera no solo se consideran las bases y almacenes de datos, sino también las herramientas que permiten la generación de modelos como por ejemplo Weka o Python.

Debido al auge de los datos y la analítica, las instituciones de educación superior han iniciado investigaciones en este campo. Así pues, la analítica del aprendizaje surge dentro de este panorama como la medición, recolección, análisis y divulgación de datos sobre los alumnos y sus contextos, con el propósito y comprender y optimizar el aprendizaje y los ambientes en que se produce *(Avella et al., 2016)*. Así mismo la analítica académica, fortalece el cumplimiento de los indicadores de las instituciones permitiéndoles trabajar y reorganizar eficientemente sus recursos, con el fin de mejorar aspectos tales como: el rendimiento académico, disminuir la tasa de deserción e incrementar la tasa de graduación. *(Mahroeian et al., 2017)* establecieron a través de su trabajo cuales son los posibles beneficios del uso de la analítica en educación superior **(Figura 3)**.



Figura 3. Beneficios institucionales del uso de la analítica Fuente; Adaptado de (Mahroeian et

al., 2017)

Por tanto, la analítica académica y la analítica del aprendizaje son campos emergentes, en el caso de la educación colombiana que ha traído consigo la necesidad de implementar estrategias que modifiquen la cultura de utilizar los datos generados por el estudiante en los procesos de formación y aprendizaje con el fin de que sean analizados para determinar su influencia de diversas variables en la determinación de fenómenos como la deserción académica, el rendimiento académico, y/o la tasa de graduación. Este ha sido un proceso lento en la educación superior debido a la tardía adopción de la analítica para mejorar la práctica educativa, pero fundamentalmente porque existen diferentes percepciones de los actores involucrados en cuanto a que constituye la analítica y cuál es el papel que puede desempeñar al abordar los desafíos educativos (*Mahroeian et al., 2017*).

METODOLOGÍA

La metodología para realizar la revisión bibliográfica se sintetiza en: (a) recopilación de información, (b) evaluar la idoneidad de la información, (c) analizar e interpretar datos relevantes y (d) resultados.

Recopilación de la información.

La búsqueda se llevó a cabo en bases de datos tales como: Scopus, Springer Links, IEEE Explorer, Proquest durante el periodo 2015-2020 (octubre). Las palabras clave utilizadas fueron combinaciones de términos tales como: learning, learning analytics, learning analytics research, educational factors, academic analytics, academic analytics research, E-Learning, M-learning, G-learning, academic performance prediction.

Uno de los principales obstáculos que se presentaron en la recopilación de la información fue encontrar información referente a casos de estudio de aplicación de métodos relacionados con la analítica académica en educación superior (80 artículos). Solo se tuvo en cuenta 50 artículos ya que muchos de estos son artículos de revisión de literatura o son aplicados a la educación escolar.

Evaluar la idoneidad de la información

Pretende determinar las publicaciones que busquen establecer una predicción dentro del tema de búsqueda a través de diversos métodos, posteriormente a esto se clasificarán las publicaciones de diversas maneras y aspectos relevantes soportados en gráficas. Las siguientes preguntas guiaron esta revisión:

1. ¿Cuáles son los métodos para realizar analítica en educación?
2. ¿Cuáles son los campos de acción de la analítica académica en la educación superior?
3. ¿Cómo se ha realizado la socialización de los trabajos realizados en torno a la analítica académica?
4. ¿Se continúa trabajando esta temática en los últimos años

Analizar e interpretar datos relevantes

Los 59 artículos analizados, se centraron en aspectos relacionados con analítica académica:

rendimiento académico (RA), deserción académica (DA) y Graduación (GA). Los investigadores agotaron la bibliografía utilizando el procedimiento, las palabras clave y las bases de datos descritos anteriormente. Por lo tanto, esta literatura puede no incluir fuentes no disponibles a través de los criterios y bases de datos no especificadas en este documento.

RESULTADOS

Con base en la revisión de la literatura, en esta sección se proporcionan los resultados obtenidos para responder a las preguntas de investigación.

Métodos y Herramientas de la analítica en educación

Es posible aplicar analítica a un conjunto de datos mediante el análisis estadístico, la minería de datos y el aprendizaje automático. A partir de la investigación referencial se ha podido establecer los métodos usados en los trabajos sobre este tema. Hecho que puede servir como insumo para docentes, estudiantes o investigadores que quieran profundizar en este campo, tal vez evitando la abrumadora tarea de una búsqueda bibliográfica para familiarizarse con los diferentes métodos de análisis (*Tabla 2*).

Tabla 2 Métodos usados de acuerdo a cada enfoque de la Analítica Académica.

Autor	Algoritmos y/o métodos utilizados	Muestra	Mejor resultado	Software utilizado
(Echegaray-calderon & Barrios-aranibar, 2015)	Algoritmos genéticos y redes neuronales	1271	Algoritmos genéticos y redes	No específica
(Pradeep et al., 2015)	JRip, OneR, J48 y ADtree.	670	Algoritmo PRISM	Weka
(Hernández et al., 2016)	Regresión logística, árboles de decisión y red neuronal	134	Regresión logística	No específica
(Dixson et al., 2016)	Regresión Jerárquica	609	Regresión	Mplus 7.3
(Ahamed et al., 2016)	Random Forest, Naïve Bayes y Redes Neuronales	423	Redes Neuronales	No específica
(Hasbun et al., 2016)	Arboles de decisión	4840	Arboles de decisión	R free
(Devasia et al., 2016)	Naïve Bayes	700	Naïve Bayes	No
(Merchán-rubiano et al., 2017)	Árbol de decisión J48, Random Forest	7644	Árbol J48	WEKA
(Razaque et al., 2017)	Naïve Bayes	660	Naïve Bayes	No
(Stanko et al., 2017)	Regresión lineal y logística	125	ninguno	No
(QIN et al., 2017)	Random forest	250	Random forest	No
(Costa et al., 2017)	Árbol de decisión, SVM, Redes neuronales, Naive Bayes	423	SVM	WEKA
(M. K. Khalil et al., 2018)	Correlación estadística	180	Correlación estadística	SPSS
(Wheeler et al., 2018)	Regresión logística multivariada	323	Regresión logística multivariada	No específica
(Liu et al., 2018)	Serie de tiempo	325	Serie de tiempo	No
(Limsathitwong et al., 2018)	Algoritmo de bosque aleatorio	2880	Algoritmo De bosque aleatorio	RapidMiner
(Lima et al., 2018)	k-Means y C5.0	1844	k-Means y C5.0	Excel - R
(Hirokawa, S. 2018)	SVM	480	SVM	WEKA

(Wati et al., 2018)	Árbol de decisión C4.5, Naive Bayes	279	Naive Bayes	RAPIDMINE R
(Miguéis et al., 2018)	SVM, Naive Bayes, Bosque aleatorio, Bagging, Bosting	2459	Bosque aleatorio	RAPIDMINE R
(Burgos et al., 2018)	Redes Neuronales SVM, LOGIT_ACT	104	LOGIT_ACT	-
(Kaunang & Rotikan, 2018)	Árbol de decisión, Bosque aleatorio	300	Árbol de decisión	WEKA
(David & Anastasija, 2019)	Árbol de clasificación	582	Árbol de clasificación	R free
(S. M. Ajibade et al., 2019)	Naive Bayes, árbol de decisión, KNN y Análisis discriminante	500	K-Vecinos más cercano (KNN)	Kalboard 360
(Nabizadeh et al., 2019)	Modelado de ecuaciones estructurales (SEM)	380	(SEM)	AMOS
(Martín, 2019)	Regresión lineal múltiple	508	Regresión lineal	No específica
(Kauffman et al., 2019)	Correlación estadística	77	Correlación	No específica
(Kwankajonwong et al., 2019)	Correlación estadística	137	Correlación estadística	IBM SPSS Statistics
(Herodotou et al., 2019)	Naive Bayes, árbol de clasificación (CART) y KNN	1384	Análisis KNN y CART	No Especifica
(Francis & Babu, 2019)	Algoritmo Ensemble	-	Algoritmo Ensemble	No específica
(Vora & Rajamani, 2019)	Algoritmo de clasificación híbrido	456	Algoritmo de clasificación híbrido	Java
(Salal et al., 2019)	Naive Bayes, árbol J48, Bosque aleatorio	21	Árbol de decisión J48	WEKA
(Amazona & Hernandez, 2019)	Naive Bayes, árbol de decisión, Deep Learning	300	Redes Deep Learning	RAPIDMINE R
(Qowidho et al., 2019)	Árbol de decisión ID3, C4.5	230	C4,5	RAPIDMINE R
(S. S. M. Ajibade et al., 2019)	Arbol de decision, KNN, SVM, Bagging, Bosting	500	Árbol de decisión	-
(Mohammadi et al., 2019)	KNN, Naive Bayes, árbol de decisión	230	KNN	WEKA
(Wiyono & Tegal, 2019)	KNN, árbol decisión, SVM	382	SVM	WEKA
(Adekitan & Noma-Osaghae, 2019)	Bosque aleatorio, árbol de decisión, Naive Bayes, Regresión Logística	1445	Regresión Logística	KNIME, ORANGE
(Nieto et al., 2019)	Árbol de decisión, regresión logística, bosque aleatorio	6100	Bosque aleatorio	PYTHON
(Francis & Babu, 2019)	Árbol de decisión, Naive Bayes, Redes neuronales,	-	Árbol de decisión	-
(Hussain, Zhu, Zhang, Abidi, & Ali, 2019)	Redes neuronales, regresión logística, Naive Bayes, SVM, árbol de decisión	100	SVM	MATLAB, SPSS
(Lasarte et al., 2020)	Bootstrap	1468	Bootstrap	AMOS 24
(Tan et al., 2020)	Regresión lineal	143	Regresión lineal 0,57	No específica
(Tenpipat & Akkarajitsakul, 2020)	árbol de decisión, bosque aleatorio y aumento de gradiente.	13714	árbol de decisión	No específica
(Medina et al., 2020)	Redes neuronales	500	Redes	SPSS

(Mavani, et al., 2020)	Naive Bayes	200	Naive Bayes	WEKA, R, PYTHON
(Mengash, 2020a)	árbol de decisión, SVM, Naive Bayes	2039	Redes neuronales	WEKA
(Tomasevic et al., 2020)	KNN, SVM, Redes neuronales, Árbol de decisión, Naive Bayes, regresión logística	32593	Redes neuronales	MATLAB
(Bhutto et al., 2020)	Regresión logística y SVM	480	SVM 78%	WEKA
(L. Contreras et al., 2020)	Árbol de decison, KNN, SVM, perceptrón	1620	Perceptrón	PYTHON
(Viloria et al., 2020)	ADTree, J48 RandomTree, REPTree SimpleCart	1200	AD tree 98.6%	WEKA
(Urteaga et al., 2020)	KNN, redes neuronales	-	redes neuronales 97%	R
(Castrillón et al., 2020)	J48	460	91,67%	WEKA
(Mengash, 2020b)	Redes neUronales, SVM, árbol de decisión, Naive Bayes	2.039	Redes neuronales 79.22%	WEKA
(L. E. Contreras et al., 2020)	K vecino más cercano, SVC, perceptrón y árbol de decisión	2114	Árbol de decisión 85,80%	PYTHON
(Hasan et al., 2020)	arbol de decision, bosque aleatorio, Boosting, AdaBoost, Gradient Boosting (XGBoost)	4413	AdaBoost	PYTHON
(Hassan et al., 2020)	Bosque aleatorio, árbol de decisión, Naive Bayes, Regresión Logística, KNN, SVM	772	Bosque aleatorio	PYTHON
(Zacharias & Athanasios, 2020)	Red neuronal	76	Red neuronal	WEKA

Campos de aplicación de la analítica en educación superior.

Los artículos fueron organizados en tres grandes categorías: analítica académica (Tabla 3). Dentro de cada uno de los campos se realizó una división atendiendo el área de investigación donde se usa la analítica, a saber: A continuación, se presenta una tabla clasificada por autores y años de publicación, en la cual se muestran resultados de investigación con respecto a analítica académica para el periodo 2015 -2020 en los cuales su objetivo principal fue, predecir y/o estudiar indicadores institucionales como: rendimiento académico (RA), deserción académica (DA) y la graduación de los estudiantes (GR). A partir de los resultados, se puede apreciar que la analítica ha entrado fuertemente al campo educativo en los últimos años para intentar explicar diversos indicadores de tipo académico en la educación superior. Así mismo de los tres indicadores que generalmente son usados para determinar la eficiencia del proceso educativos, la analítica académica muestra mayor tendencia a la explicación del rendimiento académico.

Tabla 3 Clasificación de trabajos según enfoque y campos de aplicación para la Analítica Académica

Graduación	(M. K. Khalil et al., 2018); (Wheeler et al., 2018);(David & Anastasija, 2019); (Vora & Rajamani, 2019); (Kauffman et al., 2019)
Deserción	(Hernández et al., 2016); (Hasbun et al., 2016); (Devasia et al., 2016); (Merchán-rubiano et al., 2017); (Liu et al., 2018); (Limsathitwong et al., 2018); (Lima et al., 2018); (Lasarte et al., 2020); (Tenpipat & Akkarajitsakul, 2020); (Medina et al., 2020), (L. Contreras et al., 2020)
Rendimiento académico	(Echegaray-calderon & Barrios-aranibar, 2015); (Pradeep et al., 2015); (Dixson et al., 2016); (Ahamed et al., 2016); (Razaque et al., 2017); (Stanko et al., 2017); (QIN et al., 2017); (Costa et al., 2017); (Hirokawa, S. 2018); (Wati et al., 2018); (Miguéis et al., 2018); (Burgos et al., 2018); (Kaunang & Rotikan, 2018); (S. M. Ajibade et al., 2019); (Nabizadeh et al., 2019); (Martín, 2019); (Kwankajonwong et al., 2019); (Herodotou et al., 2019); (Francis & Babu, 2019); (Vora & Rajamani, 2019); (Salal et al., 2019); (Amazona & Hernandez, 2019); (Qowidho et al., 2019); (S. S. M. Ajibade et al., 2019); (Mohammadi et al., 2019); (Wiyono, Abidin, Bersama, Lor, & Tegal, 2019); (Adekitan & Noma-Osaghae, 2019); (Nieto et al., 2019); (Francis & Babu, 2019); (Hussain, Zhu, Zhang, Abidi, & Ali, 2019); (Tan et al., 2020); (Mavani, et al., 2020); (Mengash, 2020a); (Tomasevic et al., 2020); (Bhutto et al., 2020); (Viloria et al., 2020); (Urteaga et al., 2020); (Castrillón et al., 2020); (Mengash, 2020b); (L. E. Contreras et al., 2020); (Hasan et al., 2020); (Hassan et al., 2020); (Zacharias & Athanasios, 2020)

Socialización de la analítica en educación superior.

Con el fin de identificar si los trabajos relacionados con analítica académica han sido publicados por medio de revistas revisadas por pares o por medio de conferencias, se muestra en la Figura 4 un comparativo de trabajo publicados durante el periodo de investigación en revistas de alto impacto y conferencias para cada uno de los campos de la analítica aplicada en educación. Se aprecia una fuerte tendencia de los investigadores a presentar los resultados de investigaciones en revistas indexadas de alto impacto y no tanto en congresos o simposios internacionales.

Esto en nuestro criterio es debido a que la analítica es un campo en desarrollo e inclusión para la solución de problemas de tipo académico en las instituciones (Lasarte et al., 2020).

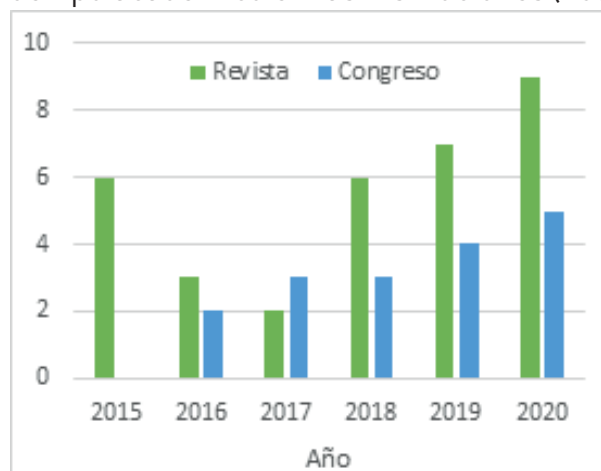


Figura 4. Medios de socialización del conocimiento discriminados por tipo de evento para AA. Cantidad de documentos que soportan la socialización del conocimiento.

En esta sección se presentan los resultados obtenidos a partir de la revisión realizada sobre los documentos consultados para el caso de Analítica académica. Los primeros años de la consulta (2015 - 2017) el número de publicaciones para Analítica académica son pocos. Pero se muestra interés por la comunidad investigativa en aplicar el uso de la analítica en diversos campos mostrados (Figura 5). Tanto así, que en los dos últimos años se nota un marcado crecimiento de investigaciones hacia la determinación del rendimiento académico en educación superior, así como para la determinación de las tasas de deserción académica y graduación.

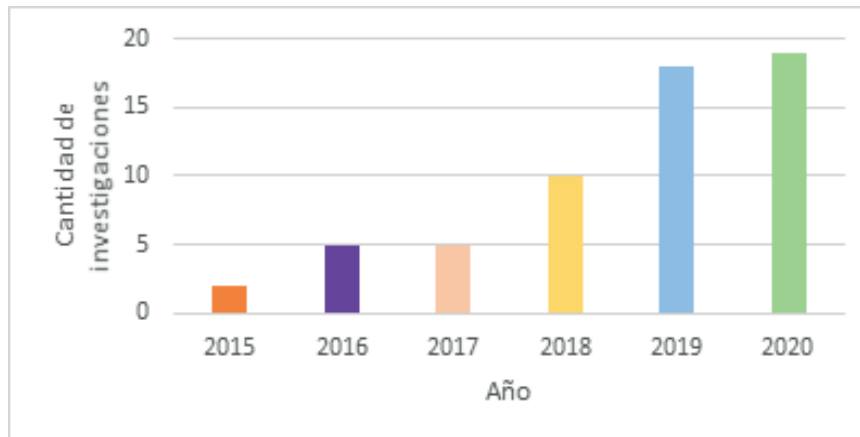


Figura 5. Número de documentos según la investigación referencial para la analítica académica en educación superior

Realizando un análisis sobre cuáles son los métodos o algoritmos que se han estado utilizando para estudiar y/o predecir algunos aspectos dentro del campo de la analítica académica, se elaboró la Figura 6 donde se aprecia que las técnicas estadísticas están dejando de usarse en investigaciones aplicadas a la educación cuando de predicción de fenómenos se trata. El enfoque hacia el campo de la minería de datos y el aprendizaje automático se evidencia con el uso de algoritmos como redes neuronales, arboles de decisión y máquinas de vectores de soporte (SVM).

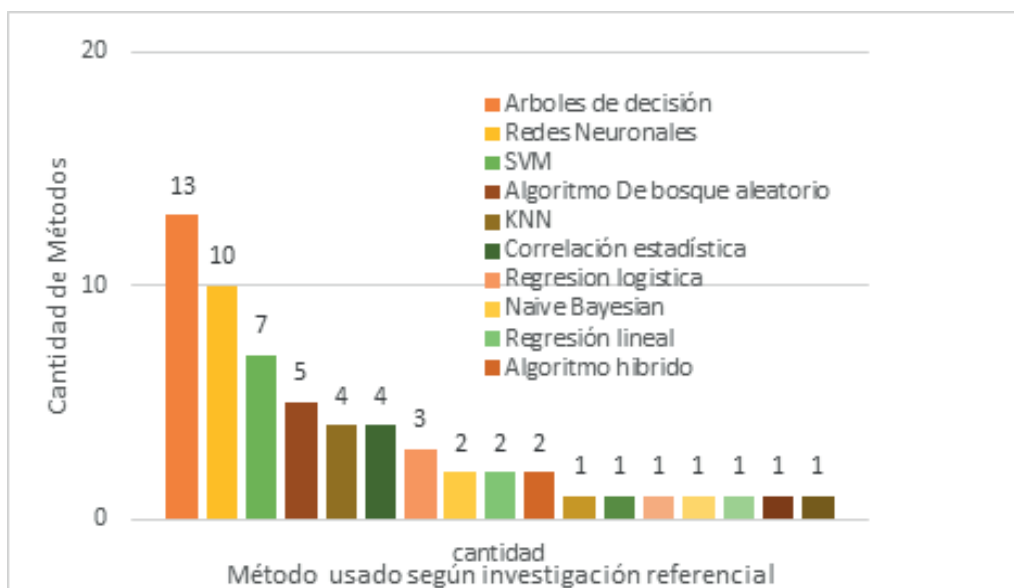


Figura 6. Algoritmos o métodos usados en trabajos relacionados con analítica Académica a partir de la revisión bibliográfica

Productividad de los países.

Realizando un análisis sobre los países que están trabajando cada una de las temáticas relacionadas con analítica académica, y han producido documentos relacionados, en la **Figura 7** se muestra que la India es el país con el que mayor frecuencia se comparte autoría. De los documentos obtenidos, un alto porcentaje aprox. (65%) son de documentos desarrollados en Asia y Europa para el caso de analítica académica, seguidos de América del sur.

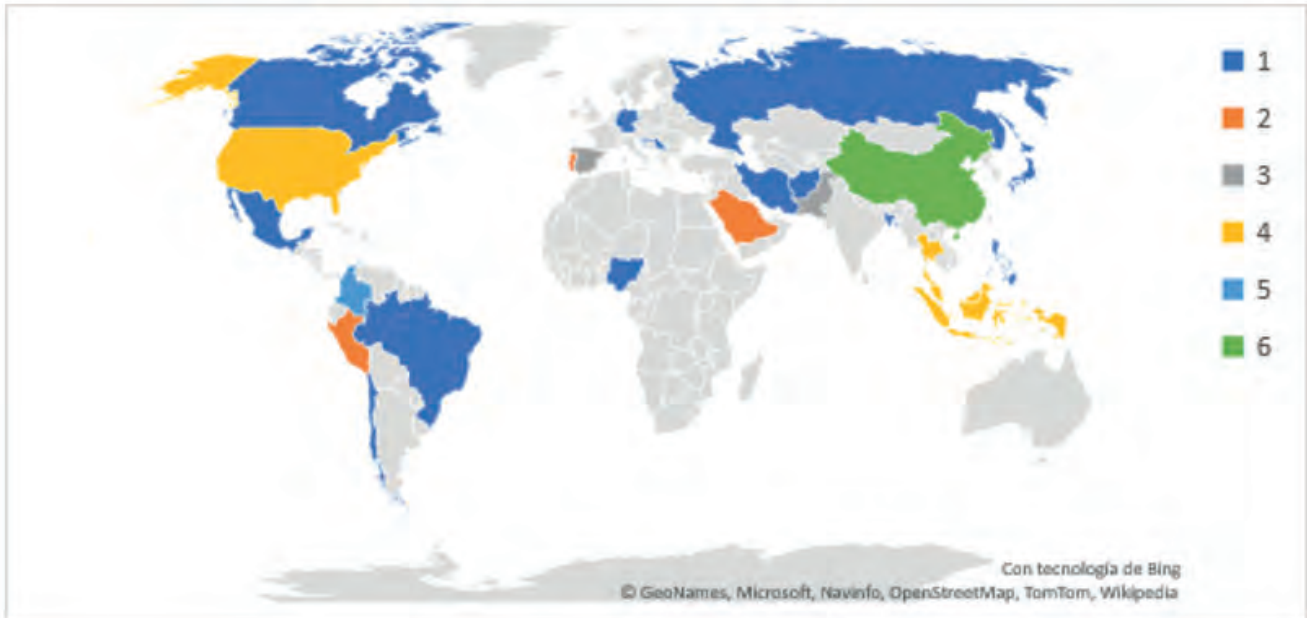


Figura 7. Países que han realizado más publicaciones de analítica académica según la revisión bibliográfica.

CONCLUSIONES

De acuerdo al trabajo presentado y a los resultados obtenidos, se pueden plantear las siguientes conclusiones principales:

- 1.- Los investigadores se han preocupado en los últimos años a trabajar en el desarrollo de modelos que permitan comprender aspectos de la vida académica del estudiante, docentes e instituciones que permitan la elaboración y toma de decisiones acertadas.
- 2.- La analítica es un campo en crecimiento usado dentro del ámbito educativo que tiene por objetivo, estudiar cómo influyen los factores tecnológicos y de gestión en las instituciones, por ejemplo, aspectos relacionados con la organización y su área financiera. (Oblinger et al., 2007). Busca la aplicación de los principios y herramientas de la minería de datos y aprendizaje automático a la academia.
- 3.- El rendimiento académico, la tasa de deserción y la tasa de graduación en su respectivo orden constituyen factores imprescindibles en el abordaje del tema de la calidad de la educación superior, siendo el primero el más influyente según las investigaciones analizadas.
- 4.- El uso de herramientas tecnológicas avanzadas dentro de las ciencias de la computación como lo es el aprendizaje automático y sus algoritmos están apoyando y fortaleciendo la toma de decisiones desde el punto de vista operacional, académico y financiero.
- 5.- De acuerdo con los encontrado en el análisis de la investigación referencial, es posible utilizar los algoritmos de aprendizaje automático en un enfoque para estimar la cantidad de

estudiantes graduados por cohorte y los mejores campos de acción de un grupo de estudiantes al terminar la carrera universitaria, así como a los indicadores anteriormente mencionados.

Pero tal como lo menciona (Daniel, 2015a), la tecnología para entregar este potencial es aún muy joven y la investigación en la comprensión de la utilidad pedagógica está todavía en su infancia.

REFERENCIAS BIBLIOGRÁFICAS

Adekitan, A. I., & Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, 24(2), 1527–1543. <https://doi.org/10.1007/s10639-018-9839-7>

Ahamed, A. T. M. S., Mahmood, N. T., & Rahman, R. M. (2016). Prediction of HSC Examination Performance using Socioeconomic , Psychological and Academic Factors. *International Conference on Electrical and Computer Engineering*, 263–266.

Ajibade, S. M., Ahmad, N. B., & Shamsuddin, S. M. (2019). An Heuristic Feature Selection Algorithm to Evaluate Academic Performance of Students. *10th Control and System Graduate Research Colloquium (ICSGRC)*, 2–3.

Ajibade, S. S. M., Bahiah Binti Ahmad, N., & Mariyam Shamsuddin, S. (2019). Educational Data Mining: Enhancement of Student Performance model using Ensemble Methods. *IOP Conference Series: Materials Science and Engineering*, 551(1). <https://doi.org/10.1088/1757-899X/551/1/012061>

Amazona, M. V., & Hernandez, A. A. (2019). Modelling Student Performance Using Data Mining Techniques. *Proceedings of the 2019 5th International Conference on Computing and Data Engineering - ICCDE' 19*, 36–40. <https://doi.org/10.1145/3330530.3330544>

Argonza, J. (2016). Big data en la educación. *Revista Digital Universitaria*, 17(1), 1–16. <http://www.revista.unam.mx/vol.17/num1/art06/>

Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Journal of Asynchronous Learning Network*, 20(2). <https://doi.org/10.24059/olj.v20i2.790>

Baepler, P., & Murdoch, C. (2010a). Academic Analytics and Data Mining in Higher Education. *International Journal for the Scholarship of Teaching and Learning*, 4(2). <https://doi.org/10.20429/ijstl.2010.040217>

Baepler, P., & Murdoch, C. J. (2010b). International Journal for the Scholarship of Teaching and Learning Academic Analytics and Data Mining in Higher Education Academic Analytics and Data Mining in Higher Education. *International Journal for the Scholarship of Teaching and Learning*, 4(2). <https://doi.org/10.20429/ijstl.2010.040217>

Bhutto, S., Siddiqui, I. F., Arain, Q. A., & Anwar, M. (2020). Predicting Students' Academic Performance Through Supervised Machine Learning. *ICISCT 2020 - 2nd International Conference on Information Science and Communication Technology*. <https://doi.org/10.1109/ICISCT49550.2020.9080033>

Brownlee, J. (2020). *Machine Learning Mastery*. <https://machinelearningmastery.com/>

- Burgos, C., Campanario, M. L., Peña, D. de la, Lara, J. A., Lizcano, D., & Martínez, M. A. (2018). Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers and Electrical Engineering*, 66, 541–556. <https://doi.org/10.1016/j.compeleceng.2017.03.005>
- Burman, I., & Som, S. (2019). Predicting Students Academic Performance Using Support Vector Machine. *Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019*, 756–759. <https://doi.org/10.1109/AICAI.2019.8701260>
- Campbell, J. P., & Oblinger, D. (2007). Learning analytics | Higher Education Academy. <https://www.heacademy.ac.uk/enhancement/starter-tools/learning-analytics#snapshot-logo>
- Castrillón, O. D., Sarache, W., & Ruiz-Herrera, S. (2020). Prediction of academic performance using artificial intelligence techniques. *Formacion Universitaria*, 13(1), 93–102. <https://doi.org/10.4067/S0718-50062020000100093>
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), 683–695. <https://doi.org/10.1080/13562517.2013.827653>
- Contreras, L. E., Fuentes, H. J., & Rodríguez, I. (2020). ACADEMIC INTERRUPTION MODEL USING AUTOMATIC LEARNING ALGORITHMS. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, 10(3), 16075–16086. www.tjprc.org
- Contreras, L., Fuentes, H., & Rodriguez, J. (2020). Academic performance prediction by machine learning as a success/failure indicator for engineering students. *Formacion Universitaria*, 13(5), 233–246. <https://doi.org/10.4067/S0718-50062020000500233>
- Contreras, L., & Tarazona, G. (2018). Big Data: An Exploration Toward the Improve of the Academic Performance in Higher Education. *Lecture Notes in Computer Science*, 10943, 627–637. <https://doi.org/10.1007/978-3-319-93803-5>
- Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73, 247–256. <https://doi.org/10.1016/j.chb.2017.01.047>
- Daniel, B. (2015a). Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46–5(0007–1013), 904–920. <https://www.learntechlib.org/p/151804/>
- Daniel, B. (2015b). Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46(5), 904–920. <https://doi.org/10.1111/bjet.12230>
- Daumiller, M., & Janke, S. (2019). The Impact of Performance Goals on Cheating Depends on How Performance Is Evaluated. *AERA Open*, 5(4), 1–10. <https://doi.org/10.1177/2332858419894276>
- David, J., & Anastasija, G. (2019). Predicting Academic Performance Based on Students' Family Environment : Evidence for Colombia Using Classification Trees. *Psychology, Society, & Education*, 11(3), 299–311. <https://doi.org/10.25115/psyse.v10i1.2056>
- Devasia, M. T., P, M. V. T., & Hegde, V. (2016). Prediction of Students Performance using Educational Data Mining. *International Conference on Data Mining and Advanced Computing*

(SAPIENCE), 91–95.

Dixson, D. D., Worrell, F. C., Olszewski-Kubilius, P., & Subotnik, R. F. (2016). Beyond perceived ability: the contribution of psychosocial factors to academic performance. *Annals of the New York Academy of Sciences*, 1377(1), 67–77. <https://doi.org/10.1111/nyas.13210>

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Educational Technology and Society*, 15(3), 58–76.

Echegaray-calderon, O. A., & Barrios-aranibar, D. (2015). Optimal selection of factors using Genetic Algorithms and Neural Networks for the prediction of students' academic performance. *Congreso Latinoamericano de Inteligencia Computacional*, 1–6.

Estrada, J., & Quintero, R. (2015). Bajo rendimiento académico en la Universidad Distrital Francisco José de Caldas (Universidad Distrital (ed.); Primera). Editorial UD.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning Int. J. Technology Enhanced Learning Int. J. Technology Enhanced Learning*, 4(x), 304–317. <https://doi.org/10.1504/IJTEL.2012.051816>

Fournier, H. ;, Kop, R. ;, & Sitlia, H. (2011). The Value of Learning Analytics to Networked Learning on a Personal Learning Environment. *LAK '11 Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 104–109. <https://doi.org/10.1145/2090116.2090131>

Francis, B. K., & Babu, S. S. (2019). Predicting Academic Performance of Students Using a Hybrid Data Mining Approach. *Journal of Medical Systems*, 43(6). <https://doi.org/10.1007/s10916-019-1295-4>

García-González, J. R., Sánchez-Sánchez, P. A., Orozco, M., & Obredor, S. (2019). Extracción de Conocimiento para la Predicción y Análisis de los Resultados de la Prueba de Calidad de la Educación Superior en Colombia Knowledge Capture for the Prediction and Analysis of Results of the Quality Test of Higher Education in Colombia. *Revista Formación Universitaria*, 12(4), 55–62. <https://doi.org/10.4067/S0718-50062019000400055>

García Tinisaray, D. K. (2015). CONSTRUCCIÓN DE UN MODELO PARA DETERMINAR EL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES BASADO EN LEARNING ANALYTICS (ANÁLISIS DEL APRENDIZAJE), MEDIANTE EL USO DE TÉCNICAS MULTIVARIANTES. Universidad de Sevilla.

Goldstein, P. J., & Katz, R. N. (2005). Academic Analytics : The Uses Of Management Information And Technology In Higher Education. *EDUCAUSE Quarterly*, 8(December), 113. <https://library.educause.edu/resources/2005/12/academic-analytics-the-uses-of-management-information-and-technology-in-higher-education>

Gómez, D. (2015). Analítica visual en eLearning [Universdiad de Salamanca]. <https://gredos.usal.es/handle/10366/125950>

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology and Society*, 15(3), 42–57.

Hasan, R., Palaniappan, S., Mahmood, S., Abbas, A., Sarker, K. U., & Sattar, M. U. (2020).

Predicting student performance in higher educational institutions using video learning analytics and data mining techniques. *Applied Sciences*, 10(11). <https://doi.org/10.3390/app10113894>

Hasbun, T., Araya, A., & Villalon, J. (2016). Extracurricular activities as dropout prediction factors in higher education using decision trees. 16th International Conference on Advanced Learning Technologies Extracurricular, 3–5. <https://doi.org/10.1109/ICALT.2016.66>

Hassan, H., Ahmad, N. B., & Anuar, S. (2020). Improved students' performance prediction for multi-class imbalanced problems using hybrid and ensemble approach in educational data mining. *Journal of Physics*, 1529(5). <https://doi.org/10.1088/1742-6596/1529/5/052041>

Hernández, A. G., Meléndez, R. A., Morales, L. A., García, A., Tecpanecatl, J. L., & Algreto, I. (2016). Comparative Study of Algorithms to Predict the Desertion in the Students at the ITSM-Mexico. *LATIN AMERICA TRANSACTIONS*, 14(11), 4573–4578.

Herodotou, C., Rienties, B., Boroowa, A., & Zdrahal, Z. (2019). A large - scale implementation of predictive learning analytics in higher education : the teachers ' role and perspective. In *Educational Technology Research and Development* (Vol. 67, Issue 5). Springer US. <https://doi.org/10.1007/s11423-019-09685-0>

Hirokawa, S. (2018). Key attribute for predicting student academic performance. *ACM International Conference Proceeding Series*, 308–313. <https://doi.org/10.1145/3290511.3290576>
Iten, L., Arnold, K., & Pistilli, M. (2008). Mining Real-Time Data to Improve Student Success in a Gateway Course. Eleventh Annual TLT Conference.

Janani, P., Verma, S., Natarajan, S., & Sinha, A. K. (2019). Communication Using IoT. In *Information and Communication Technology for Sustainable Development, Advances in Intelligent Systems and Computing*. Springer Singapore. <https://doi.org/10.1007/978-981-13-7166-0>

Kauffman, C., Derazin, M., Asmar, A., & Kibble, J. (2019). Patterns of medical student engagement in a second-year pathophysiology course: relationship to USMLE Step 1 performance. *Advances in Physiology Education*, 43(01), 512–518. <https://doi.org/10.1152/advan.00082.2019>

Kaunang, F. J., & Rotikan, R. (2018). Students' academic performance prediction using data mining. *Proceedings of the 3rd International Conference on Informatics and Computing, ICIC 2018*, 1–5. <https://doi.org/10.1109/IAC.2018.8780547>

Khalil, M. K., Hawkins, H. G., Crespo, L. M., & Buggy, J. (2018). The Design and Development of Prediction Models for Maximizing Students ' Academic Achievement. *Medical Science Educator*, 111–117.

Khalil, M., Khalil, M., & Ebner, M. (2015). Learning Analytics: Principles and Constraints. *EdMedia: World Conference on Educational Media and Technology*, 2015(1), 1789–1799. <https://www.learntechlib.org/p/151455/>

Kwankajonwong, N., Ongprakobkul, C., Qureshi, S. P., Watanatada, P., Thanprasertsuk, S., & Bongsebandhu-phubhakdi, S. (2019). Attitude, but Not Self-Evaluated Knowledge, Correlates with Academic Performance in Physiology in Thai Medical Students. *Advances in Physiology Education*, 43(3), p324-331.

Lasarte, O. F., Díaz, E. R., Palacios, E. G., & Fernández, A. R. (2020). The role of social support in school adjustment during Secondary Education. 32(1), 100–107. <https://doi.org/10.7334/>