

Big Data, una estrategia para evitar la deserción escolar en las IES

Arturo Amaya-Amaya, Franklin Huerta-Castro y Carlos O. Flores-Rodríguez

RESUMEN

La diversificación de modalidades educativas como *e-Learning* y *b-learning* ha permitido aumentar los índices de cobertura de la educación superior. Pero en la medida en que más estudiantes ingresan a las universidades, son más los estudiantes que no logran finalizar sus estudios de pregrado, detonando con ello los índices de deserción escolar, problema que no sólo tiene altos costos económicos y sociales a nivel nacional e internacional, sino que también genera condiciones de exclusión y de pobreza. El siguiente trabajo analiza los efectos negativos emanados de la deserción escolar universitaria, así como las características del *Big Data*, siendo éste una solución tecnológica viable y pertinente para brindar respuestas a esta problemática. Por otra parte, también se presentan las características de la implementación del Modelo Analítico de *Big Data* de la Universidad Autónoma de Tamaulipas (UAT), así como sus resultados que permitieron identificar causas y factores que inciden en la deserción de los estudiantes universitarios.

Palabras clave: *Big Data*, modelo analítico, deserción escolar, México.

Arturo Amaya-Amaya

arturo.amaya@docentes.uat.edu.mx

Mexicano. Doctorado en Educación Internacional, Universidad Autónoma de Tamaulipas (UAT), México; Maestría en *Big Data & Business Intelligence*, Universidad de Zaragoza, España; Maestría en Administración de Empresas, Universidad Autónoma de Nuevo León, México. Profesor Investigador de la UAT, México. Temas de investigación: *Big Data*, educación a distancia.

Franklin Huerta-Castro

fhuerta@uat.edu.mx

Mexicano. Maestría en *Big Data & Business Intelligence*, Universidad de Zaragoza, España. Profesor Investigador de la Universidad Autónoma de Tamaulipas, México. Temas de investigación: *Big Data*.

Carlos O. Flores-Rodríguez

cofloresrdz@gmail.com

Mexicano. Maestría en *Big Data & Business Intelligence*, Universidad de Zaragoza, España. Profesor Investigador de Empresa ExchangeHub LLC – Atlanta, GA. Temas de investigación: *Big Data*.

Big data, uma estratégia para evitar o abandono escolar nas IES

RESUMO

A diversificação de modalidades educativas como *e-learning* e *b-learning* têm permitido aumentar os índices de cobertura da educação superior. Mas na medida em que mais estudantes entram nas universidades, são mais os que não conseguem finalizar seus estudos de graduação, detonando com isso os índices de abandono escolar, problema que não só tem altos custos económicos e sociais a nível nacional e internacional, mas que também gera condiciones de exclusão e de pobreza. O seguinte trabalho analisa os efeitos negativos emanados do abandono escolar universitário, assim como as características do *Big Data*, sendo este uma solução tecnológica viável e pertinente para brindar respostas a esta problemática. Por outro lado, também se apresentam as características da implementação do Modelo Analítico de *Big Data* da Universidade Autónoma de Tamaulipas (UAT), assim como seus resultados, que permitiram identificar causas e fatores que incidem no abandono dos estudantes universitários.

Palavras chave: *Big Data*, modelo analítico, abandono escolar, México.

Big data, a strategy to prevent academic dropout in HEIs

ABSTRACT

The diversification of educational modalities such as *e-learning* and *b-learning* has made it possible to increase coverage rates in higher education. However, as more students enter universities, more of them fail to complete their undergraduate studies, thus detonating dropout rates, a problem that not only has high economic and social costs at the national and international level, but also generates conditions of exclusion and poverty. The following article analyzes the negative effects of academic dropout, as well as the characteristics of *Big Data*, which is a viable and pertinent technological solution to provide answers to this problem. On the other hand, the authors also present the main characteristics of the implementation of the *Big Data* Analytical Model of the Universidad Autónoma de Tamaulipas (UAT), as well as its results, which allowed to identify causes and factors that influence university students' desertion.

Key words: *Big Data*, analytical model, dropout, Mexico.

Recepción: 12/04/18. **Aprobación:** 06/05/19.



Introducción

La deserción escolar en la educación superior muestra en diferentes partes del mundo cifras que han merecido una atención especial tanto a nivel de las autoridades políticas, como de los líderes académicos y de la sociedad en general. En Europa, el denominado Espacio Europeo de Educación Superior (EEES), cuya creación se dio como fruto del proceso de Bolonia, se ha dado como prioridad que los 47 países miembros establezcan mecanismos para que los estudiantes puedan completar sus estudios sin obstáculos en relación con su entorno social y económico. La deserción en la mayoría de los países que conforman el EEES oscila entre el 20% y el 55%, lo que se ha acentuado por los grandes procesos migratorios y la crisis financiera de finales de la década pasada. La expansión de la educación superior en Europa no ha beneficiado a todos los grupos de la misma manera y por lo tanto, los grupos socioeconómicos más vulnerables son más propensos a la deserción escolar (Quinn, 2013).

En Estados Unidos, donde más se ha estudiado el problema, se estima que unos 400 000 estudiantes abandonan la universidad cada año. Más del 40% de los estudiantes estadounidenses que comienzan programas de cuatro años no obtienen su grado en seis años. En los programas de dos años, realizados en los llamados *Community College*, los resultados son aún peores. Lo más relevante es que el fenómeno es más dramático cuando se estudia la deserción vinculada al nivel socioeconómico de los estudiantes, medidos por los niveles de ingresos de la familia. Allí se observa que un estudiante de una familia cuyo ingreso está en más de US\$ 90 000 anuales, tienen una probabilidad de uno entre dos de graduarse de un pregrado a los 24 años de edad, mientras para los jóvenes que pertenecen a familias cuyos ingresos son menores a US\$ 35,000 la probabilidad se reduce a uno entre 17 (*The Chronicle of Higher Education*, 2015).

A partir de datos sobre la juventud en varios países latinoamericanos, se señala que las diferentes condiciones socioeconómicas interactúan con distintos

niveles de calidad de las universidades, produciendo resultados de inequidad. “Desde estas perspectivas, la pobreza y la desigualdad en oportunidades educativas parecen jugar un papel mucho más importante sobre la deserción escolar que los mismos factores institucionales” (De Vries *et al.*, 2011: 35).

En América Latina, donde las desigualdades sociales son tan marcadas, las tasas de deserción son muy altas, oscilan entre el 40% y el 70% a lo largo de la región, y contribuyen a ampliar estas desigualdades. Con frecuencia son los estudiantes provenientes de los grupos de población más vulnerables aquellos que no logran terminar sus estudios (Salmi, 2016: 3).

México no es la excepción, tal es el caso que en el Plan Sectorial de Educación 2013-2018 del Gobierno Federal de los Estados Unidos Mexicanos, en su Capítulo 1, relacionado con el Diagnóstico de la Educación Superior y Formación para el Trabajo, señala que el abandono de los estudios en la educación superior es, sin duda, uno de los desafíos más severos del sistema educativo.

La deserción escolar se presenta como un problema multifactorial con diferentes aristas que no se puede atender con el análisis de información limitada. En este sentido, no existe una receta de cocina para poder identificar con precisión las causas y factores que detonan la deserción escolar universitaria, debido a que en cada contexto educativo las condiciones y circunstancias son cambiantes y esta variabilidad depende principalmente de aspectos socioeducativos y de la realidad social en la cual están inmersos los estudiantes.

De acuerdo con la Organización para la Cooperación y el Desarrollo Económicos (OCDE), en México solo el 38% de los jóvenes que cursan la universidad logran graduarse. Por su parte, la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) menciona que “el promedio nacional de Eficiencia Terminal (ET) de la educación superior, considerando titulación, es del 39%” (López *et al.*, 2008: 136).

La deserción escolar

El fenómeno de la deserción escolar se debe a múltiples variables que además interactúan entre sí. La organización universitaria puede crear trabas para el estudiante (Tinto, 1993), el estudiante puede tener características que le dificultan integrarse (Bean, 1990) y el involucramiento que siente el estudiante con la institución puede ser importante (De Vries *et al.*, 2011: 33).

González (2006) señala como posibles factores de deserción la falta de apoyos financieros para estudiantes e instituciones, las altas tasas de desempleo juvenil, las insuficiencias en la preparación previa, la ausencia de una orientación vocacional, así como la falta de preparación y actualización de los docentes para brindar a los estudiantes el acompañamiento, orientación y seguimiento en su quehacer académico.

Por su parte, Contreras (2014) asegura que los estudiantes que muestren una caída significativa de las notas en el curso de un semestre pueden tener un 20% de posibilidades de abandonar la escuela, y se podría suponer que los estudiantes de primera generación tienen una probabilidad del 30% de deserción escolar, pero los estudiantes en ambas categorías podrían tener un 60% o incluso más probabilidades de abandonar la universidad. Cada nuevo indicador que consideramos nos acerca a la identificación de los estudiantes que están más necesitados de asistencia y orientación, permitiéndonos prevenir su deserción.

Para el Sistema Educativo Mexicano, la deserción escolar universitaria se ha convertido en uno de los desafíos más importantes:

Este fenómeno no sólo tiene altos costes económicos y sociales, sino que perpetúa las condiciones de exclusión y de pobreza. Las causas del problema incluyen factores económicos, educativos y familiares. Su reducción se convierte en un objetivo relevante de política educativa e implica, entre otras, el impulso de acciones articuladas, orientadas a apoyar a

los jóvenes en situación de desventaja, una creciente profesionalización docente y de la gestión escolar, mayor pertinencia de los planes y programas de estudio, y la prevención de los riesgos que afectan a los jóvenes (SEP, 2013: 30).

El fenómeno de la deserción en la educación superior se ha presentado durante décadas a nivel mundial. Sin embargo, ha tenido poca atención hasta hace apenas un par de lustros cuando se empezó a trabajar en políticas educativas nacionales para ampliar la cobertura de educación superior, sin considerar que entre más estudiantes ingresan a las universidades, son más los estudiantes que no logran finalizar sus estudios de pregrado, detonando con ello los índices de deserción escolar.

Las políticas de desarrollo de la educación superior en los países se han centrado en incrementar la tasa de cobertura, en muchos casos de manera masiva, sin tener en consideración los diferentes tipos de estudiantes que ingresaban con características y niveles de preparación muy diversos (Salmi, 2016: 2).

El Big Data

Después de analizar las problemáticas sociales que implica la deserción escolar en las instituciones de educación superior (IES), así como sus altos costos, es el momento de analizar las características del *Big Data* para poder identificar y potencializar sus funcionalidades como una estrategia viable y pertinente para evitar la deserción escolar universitaria.

Según Tascón (2013), las principales características que han definido el término de *Big Data* han sido las denominadas tres “V”: volumen, variabilidad y velocidad. Por otra parte, las “V” relacionadas con veracidad, valor y visualización también hacen la diferencia en comparación con otro tipo de soluciones tecnológicas. En la figura 1 se describe de una manera clara y sencilla las seis “V” que caracterizan a las soluciones tecnológicas *Big Data*.



Figura 1. Las seis "V" del Big Data



Fuente: elaboración propia.

En *Big Data* se puede trabajar con datos estructurados y no estructurados, los primeros provienen de fuentes de información conocidas, es decir, pueden ser datos emanados de los sistemas de información de la misma institución, organización o empresa, los cuales son fáciles de medir y analizar a través de los mismos sistemas de información tradicionales. Pero el verdadero reto y potencial de *Big Data* se encuentra en la posibilidad de trabajar y analizar datos no estructurados: los que provienen de aplicaciones tecnológicas disponibles en la nube, como páginas web, redes sociales, dispositivos móviles, sensores de edificios inteligentes, los cuales están en constante crecimiento y se mueven a alta velocidad, además de presentarse en diferentes formatos y estructuras de composición, haciendo más compleja su integración para la analítica de los datos.

Según la *International Data Corporation (IDC)*, se estima que el 90% de los datos que se trabajan en *Big Data* son datos no estructurados (Jiménez, 2016).

Otros temas relacionados con el *Big Data* son la minería de datos educacional (*Educational data mining*) y el aprendizaje analítico (*Learning analytics*), en los cuales se realizan procesos de extracción, recopilación, interpretación, medición y visualización de los resultados. La utilización de la analítica en educación ha crecido gradualmente en los últimos años por cinco razones: aumento sustancial en la cantidad de datos educativos, una mejora en los estándares de los formatos de datos, avances en la programación de aplicaciones educativas, incremento en la tecnología de procesamiento y almacenamiento, e incremento de herramientas de cuantificación y visualización de los datos.

La Universidad Autónoma de Tamaulipas (UAT), al igual que muchas IES nacionales e internacionales, sufre el problema de la deserción escolar. La UAT, consciente de esta problemática, implementó su Modelo Analítico *Big Data*, el cual parte de la idea de que la deserción escolar es un problema multifactorial con diferentes aristas que no se puede atender con el análisis de información limitada. Es decir, el Modelo Analítico *Big Data* de la UAT permite cruzar e integrar datos proveniente de diferentes asociaciones, organizaciones e instituciones externas a la UAT, con datos emanados de los propios sistemas de información institucionales, con el objetivo de identificar con mayor certeza las causas y factores que inciden en la deserción escolar universitaria.

Metodología

El universo se compone de 35 038 estudiantes inscritos en la UAT, de los cuales 30 740 son estudiantes de nivel licenciatura, surge a colación mencionar que la UAT cuenta con 96 programas educativos de nivel licenciatura (UAT, 2016).

En una primera etapa, la muestra de este trabajo se acotó a la cohorte generacional 2012-3/2016-3 de

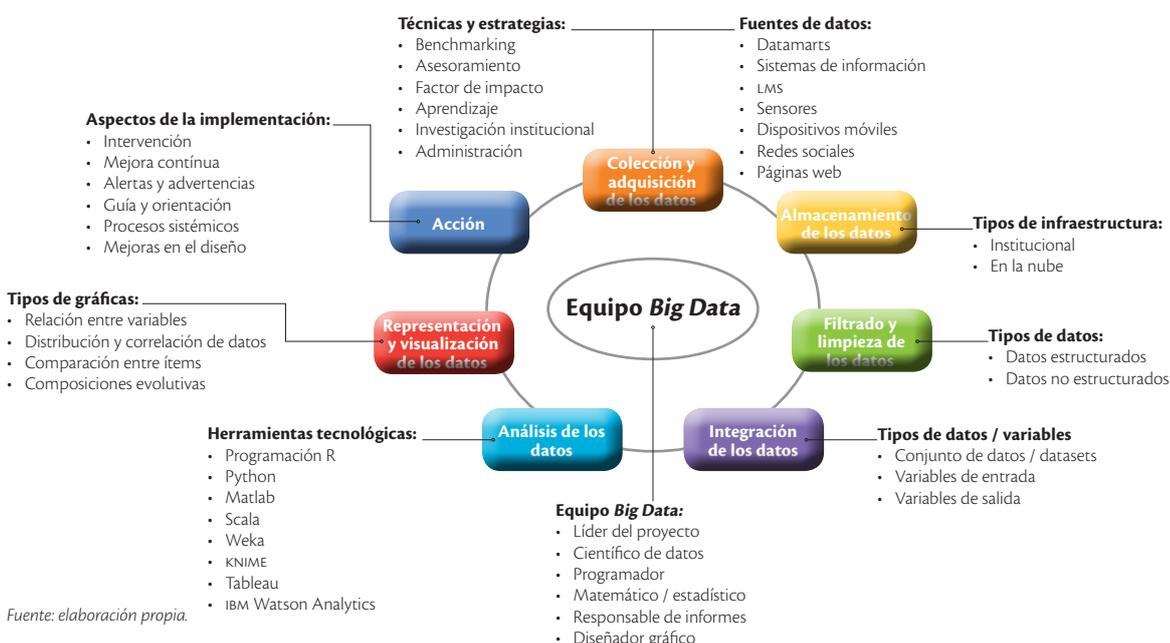
la UAT, considerando que la mayoría de los estudiantes ya egresaron, principalmente porque las licenciaturas que se ofrecen en la UAT son de cuatro años y medio. Es decir, no se consideraron las licenciaturas que tienen una duración mayor a 9 semestres, como son algunas carreras de ciencias de la salud.

En el cohorte generacional 2012-3/2016-3 se tuvo un nuevo ingreso de 8 932 estudiantes, de los cuales únicamente egresaron 3 730 estudiantes, teniendo un porcentaje de egreso del 41.75%. Es decir, 5 202 estudiantes, equivalente al 58.24% de la matrícula total, se encuentran en rezago y/o en riesgo de deserción.

Con base en lo anterior, este trabajo tiene el objetivo de identificar las posibles causas y factores que incidieron para que estos 5 202 estudiantes no egresaron en tiempo y forma, afectando con ello los índices de eficiencia terminal en las diferentes licenciaturas de la UAT.

En la figura 2 se presenta las características del Modelo Analítico *Big Data* de la UAT que se utilizó para el análisis de datos estructurados y no estructurados, emanados del Centro Nacional de Evaluación para la Educación Superior (CENEVAL) y del Sistema Integral de Información Administrativa y Académica (SIIAA) de la UAT.

Figura 2. Modelo analítico Big Data de la UAT



Fuente: elaboración propia.



Enseguida se describe cada una de las etapas del Modelo Analítico *Big Data*:

Etapa 1. Equipo *Big Data*: un científico de datos corresponde a un perfil profesional multidisciplinario, que cuenta con un nivel de estudios profesional y que tiene determinadas características, competencias y habilidades (Vázquez, 2015). Con base en lo anterior, se seleccionaron científicos de datos con experiencia en el área de la educación; experiencia en el desarrollo de soluciones tecnológicas y la informática; experiencia en el manejo de métodos y técnicas de minería de datos; así como conocimientos en las áreas de las matemáticas y la estadística para la elaboración de los modelos matemáticos.

Etapa 2. Colección y adquisición de los datos: se realizó una investigación de bases de datos internas y externas a la UAT que pudieran brindar datos relevantes. Se utilizaron bases de datos del cohorte generacional 2012-3/2016-3 relacionadas con la trayectoria escolar y la deserción escolar de los estudiantes a nivel licenciatura, emanadas del SIIAA de la UAT. También se utilizaron bases de datos externas a la UAT, como es el caso de la base de datos del Examen de Ingreso (EXANI-I) del CENEVAL.

Etapa 3. Almacenamiento de los datos: para que los datos puedan ser utilizados y analizados, deben estar almacenados físicamente en algún sistema de almacenamiento. Es decir, pueden estar almacenados en sistemas de almacenamiento institucional o en la nube a través de un centro de datos. Las bases de datos de trayectoria escolar y la deserción escolar del SIIAA de la UAT, están disponibles en sistemas de almacenamiento institucional y bajo el formato SQL (*Structured Query Language*). La base de datos del Examen de Ingreso (EXANI-I) del CENEVAL, está disponible en sistemas de almacenamiento en la nube y bajo el formato XML, en archivos SQL.

Etapa 4. Filtrado y limpieza de los datos: las bases de datos que se utilizaron en este trabajo son emanadas de sistemas de información institucionales, de tal manera que los procedimientos de limpieza

se aplicaron a datos estructurados. La limpieza de datos consiste en estandarizar la forma en que los datos van a ser ingresados a la solución tecnológica *Big Data*, muchas veces en la recolección de los datos no se siguen estos estándares, por lo cual, una de las tareas del científico de datos es preparar los datos para que se ajusten a un formato establecido. Por ejemplo, si un archivo tiene una columna llamada “Edad”, se espera que los datos en esa columna sean numéricos y que estén dentro de un rango establecido. Una celda con el dato “19” sería un dato limpio, en cambio si el dato fuera “diecinueve”, para una persona tiene el mismo sentido, pero no para la solución tecnológica *Big Data*. Sucede lo mismo si contara con datos fuera del rango, por ejemplo, números negativos, caracteres especiales o fraccionarios.

Etapa 5. Integración de los datos: las bases de datos de trayectoria escolar y la deserción escolar del SIIAA de la UAT y la base de datos del Examen de Ingreso (EXANI-I) del CENEVAL, se integraron en conjuntos de datos también conocidos como *datasets*; esto se realiza con la finalidad de tener una única fuente de datos para todo el análisis, en la cual están presentes todas las variables de entrada y de salida. Es decir, con la integración adecuada de los datos se pretende inferir cuáles son las variables de entrada que afectan directamente a las variables de salida, dicho de otra manera, cuáles son las variables que inciden en un alumno para que continúe con su carrera universitaria y cuáles son las variables que inciden en su deserción escolar.

Etapa 6. Análisis de los datos: se seleccionaron herramientas tecnológicas, métodos y técnicas de minería de datos que se utilizaron para el análisis de los datos. Existen gran cantidad de herramientas que se pueden utilizar para el análisis de datos, algunas son de software libre y otras son software propietario. Por mencionar algunas, tenemos los lenguajes de programación R, Python, Matlab y Scala; como herramientas para la minería de datos tenemos Weka y KNIME, y como herramientas para el análisis de

datos tenemos IBM Watson Analytics, Qlik View, Tableau y Microsoft PowerBI. Para efectos de la solución tecnológica *Big Data* de la UAT, se utilizaron la herramienta de minería de datos KNIME y el Método de K-Means para el análisis del clúster.

Etapa 7. Representación y visualización de los datos: se analizaron y seleccionaron aplicaciones para el diseño de diferentes tipos de gráficos que brindaron la posibilidad de presentar los datos y los resultados de la solución tecnológica *Big Data* a través de diferentes representaciones gráficas creativas e innovadoras. El ser humano, a diferencia de una máquina, no puede asimilar una gran cantidad de datos para poder tomar una decisión, pero si éstos están representados de una manera visual, el análisis de la información y resultados es más sencillo. En una gráfica podemos ver el comportamiento de nuestros datos, si van subiendo o bajando con el paso del tiempo, en qué momento fueron más altos, dónde se concentra el mayor porcentaje, etcétera. Por lo tanto, una vez que se han procesado los datos, una representación visual es muy efectiva para la correcta interpretación de los resultados.

Etapa 8. Acción: es importante mencionar que durante la implementación de la solución tecnológica *Big Data* de la UAT se programó el manejo de alertas y advertencias para notificar los estatus de los estudiantes en riesgo de reprobación y/o deserción, y poder así programar acciones de intervención con el apoyo de los actores responsables de cada uno de los procesos educativos. Por otra parte, es importante ser objetivos en los resultados emanados de la implementación de las acciones para reorientar nuestras estrategias y prevenir la deserción escolar de manera más efectiva.

Resultados

En el árbol de decisión se aprecia que la primera variable que influyó en el resultado final fue “Reprobadas en primera oportunidad”, cuando este valor es menor o igual a 0.5 existe una alta probabilidad, del 90.1%, de que el alumno alcance el estatus de Regular y Pasante. En cambio si esta variable es mayor a 0.5, la probabilidad de que el alumno alcance el estatus de suspendido temporal, baja por oportunidades de examen y baja por inscripción es del 65%. En la figura 3 se representa la interfase gráfica correspondiente con este análisis.

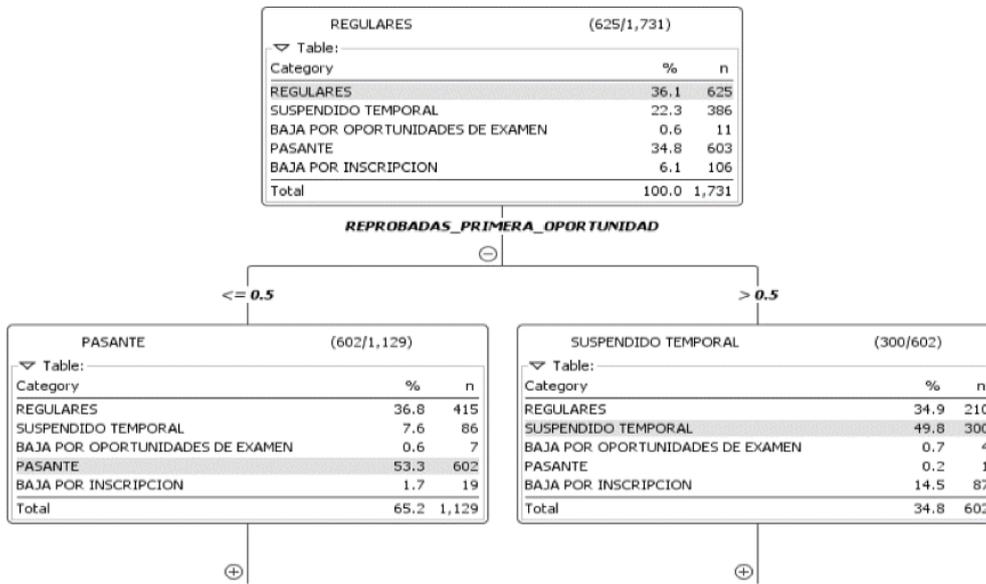
En el análisis de resultados de variables, para el caso de las mujeres, la variable más significativa fue “Vive con su padre”, en caso afirmativo las probabilidades del éxito académico son del 78%. En la figura 4 se representa gráficamente este análisis.

En el análisis de resultados de variables, para el caso de los hombres, la variable más significativa fue “Edad”, si el alumno tiene menos de 19.5 años al momento de la inscripción, las probabilidades del éxito académico se incrementan en un 13.30%, debido a que los alumnos menores a 19.5 años presentaron una eficiencia terminal del 68.1%, en comparación con los alumnos mayores a 19.5 años, los cuales presentaron un 54.8%. En la figura 5 se representa gráficamente este análisis.

En el análisis de resultados del clúster, se aprecia claramente que la única variable que diferencia un clúster de otro es la puntuación obtenida del alumno en el examen de ingreso (PCNE). Se aprecia también que el PCNE tiene una influencia en el número de asignaturas reprobadas, es decir, a mayor PCNE menor número de asignaturas reprobadas en la primera oportunidad. En la figura 6 se representa gráficamente este análisis.

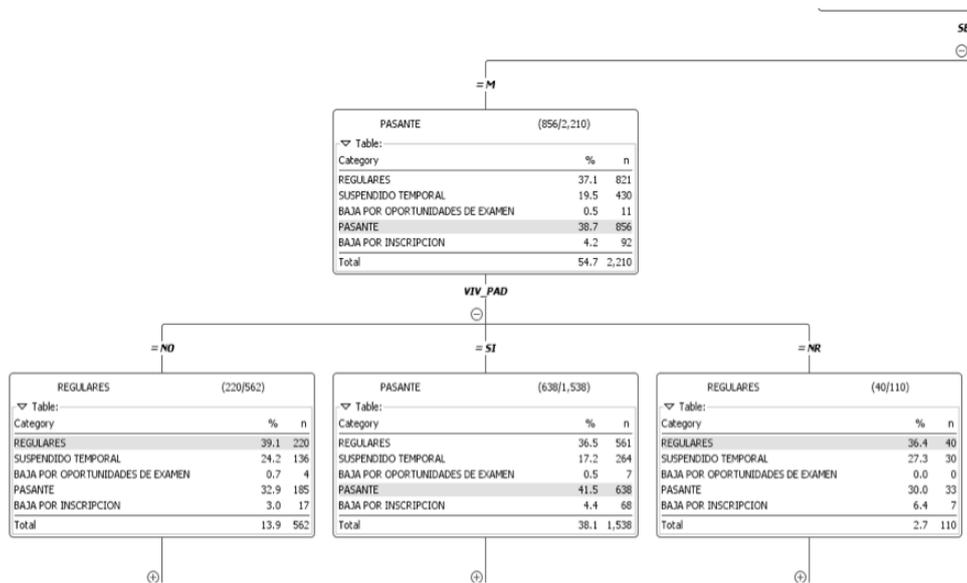


Figura 3. Análisis de resultados “Reprobadas en primera oportunidad”



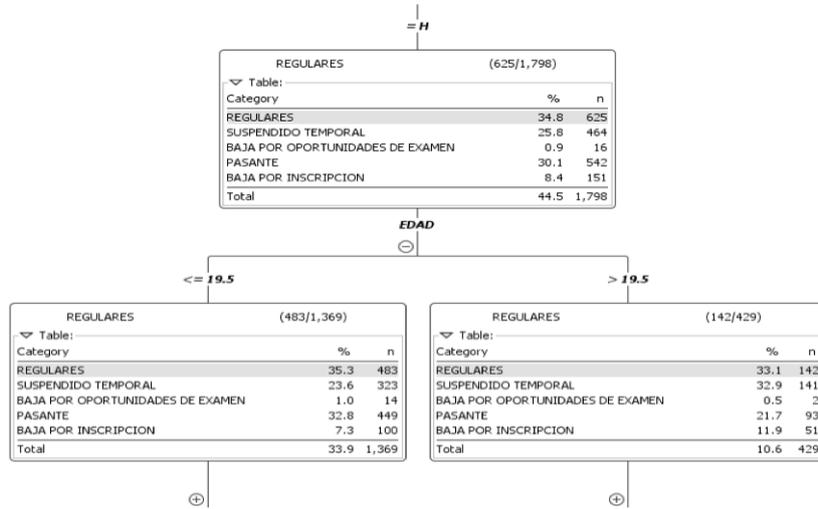
Fuente: elaboración propia.

Figura 4. Análisis de resultados “Vive con su padre”



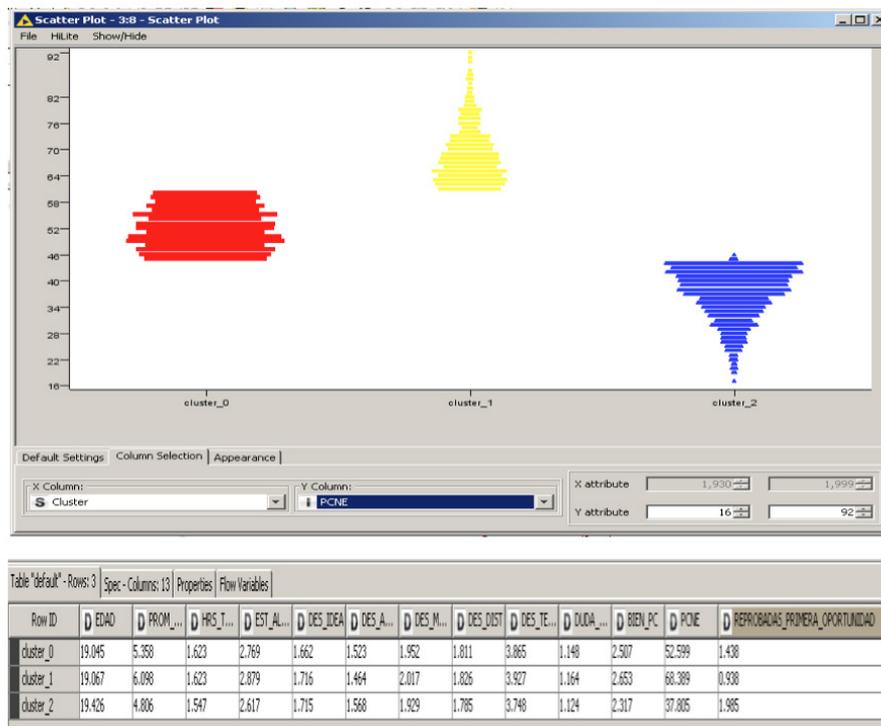
Fuente: elaboración propia.

Figura 5. Análisis de resultados “Edad”



Fuente: elaboración propia.

Figura 6. Análisis de puntuación obtenida del “Alumno en el examen de ingreso”



Fuente: elaboración propia.



Conclusiones

La retención escolar está directamente relacionada con el número de matrícula. Es decir, de ésta se derivan diferentes métricas que se utilizan en las universidades para medir cuántos estudiantes de nuevo ingreso y reingreso entran a la universidad cada inicio de periodo escolar. La matrícula es una palabra de moda desde hace mucho tiempo en los sistemas de educación superior, más aún cuando impacta en las finanzas de las IES, así como en la percepción de la sociedad. En cierto sentido, se podría decir que es cuestión de encontrar las causas y factores desencadenantes que conducen a la rotación, movilidad o deserción de los estudiantes de nivel licenciatura, pero sabemos de antemano que no es tarea fácil, tal es el caso que, hasta la fecha, los índices de deserción escolar en las IES continúan a la alza.

La UAT, convencida de que las soluciones tecnológicas *Big Data* son una estrategia viable y pertinente para revertir el problema de la deserción escolar, trabajó en esta primera etapa del proyecto, únicamente con la cohorte generacional 2012-3/2016-3; esta generación presentó un egreso del 41.75%, superior al 38% definido por la OCDE para las IES mexicanas. Para las próximas cohortes generacionales se analizarán con mayor mesura los resultados del EXANI-I del CENEVAL, debido a que se validó que a mayor PCNE menor número de asignaturas reprobadas en la primera oportunidad. Para el caso de estudiantes que obtengan una puntuación reprobatoria en el EXANI-I del CENEVAL, se considerarán como un grupo vulnerable y con alta probabilidad de deserción escolar. Con base en lo anterior, primero se identificarán las deficiencias académicas de cada estudiante y, segundo, se les dará el seguimiento a través de mecanismos de regularización para que se puedan incorporar a la licenciatura de su elección con las competencias básicas necesarias para su éxito académico.

Por otra parte, con el apoyo de la solución tecnológica *Big Data* se programarán alertas y advertencias para notificar a los actores responsables de los diferentes procesos educativos, con la intención de redoblar esfuerzos e implementar estrategias, antes y después de que los estudiantes empiecen a agotar sus oportunidades de exámenes, debido a que se validó que los estudiantes que reprueban en su primera oportunidad de examen, presentan una probabilidad del 64.95% de ser suspendidos temporalmente, tendrán baja por oportunidades de examen o baja por inscripción.

Según menciona Salmi (2016), educar a las familias sobre la importancia de acceder y finalizar la educación superior es esencial para mejorar la transición de los estudios secundarios a la educación superior. Es decir, es clave darle mucha importancia a la participación de las familias en las intervenciones, ya que es reconocido que los miembros de la familia influyen mucho en las aspiraciones de sus hijos hacia la educación superior. Tal es el caso, que en este trabajo se validó que las mujeres que viven con su padre tienen una efectividad del 78% de ser alumnos regulares y pasantes.

El *Big Data* tomará gran importancia en la educación debido a que todo apunta a que gran parte del trabajo académico de las IES se desarrollará entre el alumno e internet (Picazo, 2014), donde los sistemas de educación a distancia y las tecnologías de información (TI) disponibles en la web serán la plataforma de los procesos de enseñanza y aprendizaje, generando con ello grandes cantidades de datos para ser extraídos y analizados a través de técnicas y herramientas *Big Data*, abriendo las posibilidades de elaborar estrategias educativas con base en datos para mejorar y garantizar la efectividad del aprendizaje *ad hoc* a los ritmos y estilos de los estudiantes, ejercicio que hoy en día es difícil realizar en cualquier modalidad educativa. ■

Referencias

- Bean, J. P. (1990), *The strategic management of college enrollments*, San Francisco, Jossey-Bass.
- Contreras, C. (2014), “Deserción universitaria: abriendo los secretos de la educación con Big Data”, Centro de Investigación en Inteligencia de Negocios, <<http://www.ceine.cl/desercion-universitaria-abriendo-los-secretos-de-la-educacion-con-big-data/>> [Consulta: agosto de 2017].
- De Vries, W., P. León y J. F. Romero (2011), “¿Desertores o decepcionados? Distintas causas para abandonar los estudios universitarios”, *Revista de la Educación Superior*, vol. XL (4), pp. 29-50, <http://publicaciones.anuies.mx/pdfs/revista/Revista160_S1A3ES.pdf> [Consulta: julio de 2017].
- González, L. (2006), “Repitencia y deserción universitaria en América Latina”, en UNESCO-IESALC (2006), *Informe sobre la educación superior en América Latina y el Caribe 2000-2005: La metamorfosis de la educación superior*, <<http://es.slideshare.net/clauidiorama/unesco-iesalc-informe-sobre-la-educacin-superior-en-amrica-laytina-y-el-caribe-2005-la-metamorfosis-de-la-educacin-superior>> [Consulta: noviembre de 2017].
- Jiménez, N. (2016), “Datos no estructurados y Big Data: su importancia para las empresas”, <<https://datahoy.es/2016/05/13/datos-no-estructurados-y-big-data-su-importancia-para-las-empresas/>> [Consulta: octubre de 2017].
- López, A., A. Albiter y L. Ramírez (2008), “Eficiencia terminal en la educación superior, la necesidad de un nuevo paradigma”, *Revista de la Educación Superior*, <<http://publicaciones.anuies.mx/revista/146/5/1/es/eficiencia-terminal-en-la-educacion-superior-la-necesidad-de-un-nuevo>> [Consulta: agosto de 2017].
- Picazo, B. (2014), “La educación en 2030: una escuela menos relevante y un aprendizaje más individual”, <https://www.eldiario.es/sociedad/escuela-diferente_0_328617413.html> [Consulta: septiembre de 2017].
- Quinn, J. (2013), “Dropout and completion in higher education in Europe: among students from under-represented groups”, *An independent report authored for the European Commission*, Bruselas, European Commission, <<http://nesetweb.eu/wp-content/uploads/2015/09/2013-Drop-out-and-Completion-in-Higher-Education-in-Europe-among-students-from-under-represented-groups.pdf>> [Consulta: julio de 2017].
- Salmi, J. (2016), “Cambatir la deserción en la educación superior: lecciones a partir de la experiencia internacional”, <<http://noticias.universia.net.mx/educacion/noticia/2016/11/24/1146666/pautas-combatir-desercion-educacion-superior.html>> [Consulta: septiembre de 2017].
- Secretaría de Educación Pública (SEP) (2013), “Programa Sectorial de Educación 2013-2018”, México, SEP, <https://www.sep.gob.mx/work/models/sep1/Resource/4479/4/images/PROGRAMA_SECTORIAL_DE_EDUCACION_2013_2018_WEB.pdf> [Consulta: octubre de 2017].
- Tascón, M. (2013), “Big Data: pasado, presente y futuro”, *Revista de Pensamiento sobre Comunicación, Tecnología y Sociedad*, <<https://dialnet.unirioja.es/servlet/articulo?codigo=4423775>> [Consulta: noviembre de 2017].
- The Chronicle of Higher Education (2015), “Student success: building a culture for retention and completion on college campuses”, *Blackboard*, <<https://www.pdx.edu/president/sites/www.pdx.edu.president/files/The%20top%20trends%20and%20practices%20for%20college%20student%20retention%20and%20completion.pdf>> [Consulta: diciembre de 2017].
- Tinto, V. (1993), *Leaving college: rethinking the causes and cures of student attrition*, Chicago, The University of Chicago Press.
- Universidad Autónoma de Tamaulipas (UAT) (2016), “3er Informe Rectoral 2016”, Tamaulipas, UAT, <<https://www.uat.edu.mx/TRANS/VIIIndicadoresderesultados/Informe%20Rectoral%202016.pdf>> [Consulta: septiembre de 2017].



Vázquez, M. (2015), “El científico de datos: un perfil multidisciplinario altamente especializado”, *Revista de los Estudios de Ciencias de la Información y de la Comunicación*,

<<http://comein.uoc.edu/divulgacio/comein/es/numero46/articles/Article-Merce-Vazquez.html>> [Consulta: agosto de 2017].

Cómo citar este artículo:

Amaya-Amaya, Arturo, Franklin Huerta-Castro y Carlos O. Flores-Rodríguez (2020), “Big Data, una estrategia para evitar la deserción escolar en las IES”, *Revista Iberoamericana de Educación Superior (RIES)*, vol. XI, núm. 31, pp. 166-178, DOI: <https://doi.org/10.22201/iisue.20072872e.2020.31.712> [Consulta: fecha de última consulta].