

Deserción y rezago escolar en la Facultad de Economía de la Universidad Michoacana de San Nicolás de Hidalgo (UMSNH). Análisis a través de Minería de Datos

Dropout and School Fallback in the Faculty of Economics of Michoacana de San Nicolás de Hidalgo University (UMSNH). Analysis Through Data Mining

* Arturo García Campos

** Yiringari Pompa Flores

* Profesor de la Facultad de Economía - UMSNH.
E-mail: gcampos@umich.mx

** Profesora de la Facultad de Economía - UMSNH.
E-mail: yiringari.pompa@umich.mx

RESUMEN

La deserción escolar es un problema que se presenta no sólo en la educación primaria sino también en las instituciones de educación superior y para abatirla se han realizado esfuerzos tales como estímulos económicos y programas de acompañamiento docente para asegurar la permanencia de los estudiantes.

Este trabajo se realizó a través de minería de datos utilizando la metodología CRISP-DM para detectar los factores y variables que inciden en la deserción de los alumnos de la Facultad de Economía de la Universidad Michoacana de San Nicolás de Hidalgo (UMSNH). Los hallazgos obtenidos muestran los porcentajes de deserción y el rezago educativo; se propone una metodología para la automatización del proceso para la detección de las posibles causas de la deserción educativa.

Palabras clave: Deserción, Minería de datos, Modelo CRISP-DM, Clustering, Clasificación.

Fecha de recepción:
13 de junio de 2025

Fecha de aceptación:
17 de junio de 2025

Fecha de publicación:
19 de junio 2025

ABSTRACT

School dropout is a problem that occurs not only in primary education but also in higher education institutions. To reduce it, efforts have been made such as economic incentives and teacher support programs to ensure the permanence of students.

This work was carried out through data mining by using the CRISP-DM methodology to detect the factors and variables that affect the dropout of students from the Faculty of Economics of Michoacana de San Nicolás de Hidalgo University (UMSNH). The obtained findings show the percentage of dropouts and the educational lag; it is proposed a methodology for the automation of the process of detecting the possible causes of educational desertion.

Key words: Dropout, Data mining, CRISP-DM Model, Clustering, Classification.

INTRODUCCIÓN

Hoy en día existen carencias en el sistema educativo de nuestro país por lo que evaluar la calidad en la educación es un tema importante en todos los niveles. A lo largo del tiempo se han creado organismos no gubernamentales encargados de la calificación o valoración de los programas educativos y una de las grandes debilidades del sistema de educación radica en la incapacidad para conseguir que los estudiantes concluyan su formación ya sea en el nivel básico o universitario. Cada año hay más alumnos que se suman al rezago educativo y, por consecuencia, llegan a la deserción universitaria (Tünnermann, 1996).

Existe la preocupación en la Universidad Michoacana de San Nicolás de Hidalgo (UMSNH) de elevar o mantener la calidad en la educación, para lo cual ha establecido diferentes programas de apoyo al estudiantado como, por ejemplo, la implementación de tutorías y becas, debido a que los índices de deserción en los últimos cinco años se han mantenido anualmente en un promedio del 30% (Espinosa, 2011). En entrevista a Quadratín (2024), la UMSNH estima una deserción escolar de aproximadamente el 30% en sus

programas de licenciatura, según, el director de Control Escolar, esta cifra es más alta en las áreas de ingenierías y ciencias exactas. En contraste, las carreras con mayor retención de alumnos son las de ciencias de la salud y ciencias sociales, donde la tasa de abandono es significativamente menor, llegando a ser inferior al 10% en algunos casos como Medicina.

En ese sentido el objetivo principal de este trabajo es aplicar técnicas de minería de datos mediante la construcción y desarrollo de modelo CRISP-DM, para obtener los factores que influyen en la deserción de los alumnos de la Facultad de Economía “Vasco de Quiroga” (FevaQ) de la UMSNH. Los datos se obtuvieron de la investigación llevada a cabo por Pompa (2017), los cuales son necesarios para aplicar la técnica de minería de datos.

Este artículo se encuentra organizado en seis secciones, la primera aborda la deserción y rezago educativo. En la segunda describe la minería de datos y la investigación de la educación respecto a la problemática de la deserción estudiantil. En la tercera se desarrolla la metodología, la determinación de los objetivos y metas de la Minería de Datos. En la cuarta se implementa el modelo Scoring, mientras que en la quinta parte se exponen los resultados finales del modelado de minería de datos. Finalmente se presentan las conclusiones generales y las propuestas de la investigación.

1. Deserción y rezago educativo.

El sistema educativo en México enfrenta serias deficiencias, manifestadas en altos índices de reprobación, deserción y bajo aprovechamiento estudiantil. Las evaluaciones nacionales e internacionales han revelado que la formación actual no logra desarrollar plenamente en los alumnos habilidades cruciales para resolver problemas de forma creativa, ni los prepara adecuadamente para los desafíos de la vida y el mercado laboral. Esta problemática ha sido reconocida y plasmada en documentos oficiales como el Programa Sectorial de Educación, que se alinea con la Visión México 2030 y el Plan Nacional de Desarrollo (Presidencia de México, 2007; Villalpando, 2012).

El problema de la deserción y el rezago educativo no es nuevo en México, desde 1986, el Programa Integral para el Desarrollo de la Educación (PROIDES) ya señalaba la baja eficiencia terminal en la educación superior: en 1984 sólo el 57% de los estudiantes en 37 universidades públicas concluían sus estudios (Oca, 2007). Este diagnóstico histórico subraya la incapacidad

del sistema para que los estudiantes finalicen su formación, tanto en niveles básicos como universitarios. Aunque autores como Tinto (1989), indican algunas mejoras reportadas en 1999 en la disminución de la deserción y reprobación, la eficiencia terminal sigue siendo insatisfactoria. Comparativamente, mientras que en Estados Unidos la tasa de deserción se ha mantenido constante alrededor de 45% a lo largo del siglo XX, el fenómeno persiste en todos los niveles del sistema educativo mexicano.

La Organización para la Cooperación y Desarrollo Económicos (OCDE) ha puesto de manifiesto la alarmante trayectoria de los estudiantes mexicanos: de cada tres que ingresan a la educación media superior, poco más de la mitad la termina, solo uno de cada tres avanza a la educación superior, y apenas uno de cada seis llega a obtener una licenciatura. Estas cifras evidencian un gran desperdicio de recursos y potencial humano. Los estudios de la OCDE también indican que la deserción ocurre rápidamente, a menudo desde el primer semestre de cada ciclo escolar, siendo las causas económicas una de las principales razones (Del Castillo, 2001).

Desde hace décadas, la Secretaría de Educación Pública (SEP) y la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) han buscado soluciones al problema de la deserción en las Instituciones de Educación Superior (IES). Investigaciones y modelos de estudio se han desarrollado para entender la complejidad de la deserción y la baja eficiencia terminal (Durán y Díaz, 2001).

Sin embargo, datos recientes de la OCDE sitúan a México en el primer lugar en deserción universitaria, con un alarmante 38% de los estudiantes que ingresan a la universidad sin llegar a graduarse. Esto contrasta fuertemente con países como Países Bajos (0.07%), Finlandia (0.5%) y Alemania (4%) que presentan tasas de deserción mucho menores. Esta realidad resalta la urgencia de abordar este fenómeno (Velázquez y Romero, 2012). En la revisión sistemática realizada por Rodríguez-Camargo *et al.* (2025) a las ingenierías en México realizando la revisión de 23 artículos utilizando el criterio de exclusión e inclusión determinó que una de las principales causas de la deserción escolar en México es la económica.

La magnitud de la deserción escolar representa un problema significativo para el sistema educativo nacional, con incidencia negativa en los procesos políticos, económicos, sociales y culturales del desarrollo del país. Consciente de esto, el Plan Nacional de Desarrollo (1989-1994) ya reconocía que,

a pesar de los avances, otros factores influían negativamente en la permanencia y el rendimiento escolar. Las causas que inciden en la reprobación y deserción resultan en una baja eficiencia terminal y un desperdicio de los recursos que la sociedad invierte en educación (Durán y Díaz, 2001).

1.1. Deserción en la UMSNH

La deserción escolar en la educación básica fue uno de los temas centrales en el Primer Foro sobre el Reto de la Educación que se llevó a cabo en la UMSNH en mayo de 2013, en donde se dieron a conocer datos recientes en los que se señala que sólo el 50% de los alumnos que terminan la primaria llegan a la secundaria mientras que la cobertura en la educación media superior y superior es del 30% (Becerril, 2013). En el nivel básico las causas de la deserción se atribuyen a condiciones de marginación y dispersión poblacional en las áreas rurales y en este sentido se señaló que ampliar la cobertura y la oferta educativa en el nivel medio y superior por parte de la UMSNH no es una solución ya que el 70% de los jóvenes desertan antes de inscribirse a nivel superior (Villavicencio, 2013).

La deserción al interior de la UMSNH, es preocupante ya que los índices muestran que del 100% de los estudiantes de nuevo ingreso al bachillerato, el 60.38% desertaron sin haber concluido sus estudios, según datos obtenidos de la oficina de control escolar de la Universidad. En estudios realizados al interior de las siete preparatorias de la UMSNH se encontró que el 27% lo atribuye a problemas económicos, por otra parte, el 19% dejó la escuela por causas personales y por último el 10% fueron causas familiares.

Los índices de deserción en los últimos cinco años en las carreras que oferta la Máxima casa de Estudios de Michoacán, se han mantenido anualmente en un promedio de 30%, este fenómeno se debe por varios factores: a) alumnos que eligen una carrera diferente, lo que se traduce en una falta de orientación vocacional; b) rezago educativo (Espinosa, 2011), unas de las más importantes; c) embarazos no deseados; y d) la falta de habilidades para leer y escribir (Catalán, 2011).

Recientemente una investigación realizada por Rodríguez (2018) se basó en una encuesta de Likert dirigida a los actores del problema, obteniendo los siguientes resultados: la UMSNH constata una elevada incidencia de deserción en sus programas de licenciatura, con una mayor afectación en

las estudiantes durante los primeros semestres, sin que se perciba una tendencia a la baja en esta problemática, las causas primordiales identificadas son de índole económica y psicológica, aunque también se reconocen factores organizacionales y culturales como contribuyentes. A pesar de contar con un diagnóstico detallado del fenómeno, los funcionarios universitarios expresan que las acciones implementadas no han logrado la efectividad deseada en la reducción de la deserción. Las implicaciones de este problema se consideran trascendentales, generando pérdidas académicas tanto para el estudiante como para la institución, además de conllevar consecuencias personales como frustración y una potencial disminución en la calidad de vida del desertor, e incluso impactar en la seguridad social.

En otra investigación en la Facultad de Contaduría y Ciencias Administrativas, Meza Bucio et al. (2022) investigó las causas de deserción escolar a partir de problemas multifactoriales que afectan a la UMSNH como las huelgas, tomas, problemas políticos hasta situaciones de disminución de la matrícula de ingreso, permanencia y egreso, además, de tratar de paliar la deserción escolar de sus alumnos, encontrando que la principal causa de deserción entre sus alumnos es el relacionado a la salud.

2. Minería de datos en la educación

La minería de datos es el análisis de un conjunto de datos en observación para encontrar relaciones insospechadas y para resumir los datos de manera novedosa que sean comprensibles y útiles para el propietario (David et al., 2001). La minería de datos es un campo interdisciplinario que reúne las técnicas de aprendizaje automático, reconocimiento de patrones, estadísticas y la extracción de información de bases de datos (Cabena et al., 1998), así como la identificación de estructuras de datos, donde la estructura designa patrones estadísticos o modelos predictivos de los datos y las relaciones entre las partes de los datos (Fayyad y Uthurusamy, 2002).

En varios estudios estadísticos y de predicción realizados en diferentes instituciones educativas en México y en el mundo respecto a la deserción, se presentan diferentes variables y enfoques las cuales se desarrollan mediante diferentes métodos y modelos, la mayoría de las investigaciones toman como referencia común las variables socioeconómicas y académicas. Algunos autores presentan sólo las académicas y psicológicas, como Porcel et al. (2009) en su trabajo de modelos predictivos y técnicas de minería de datos,

proponen la construcción de modelos matemáticos que permitan predecir el rendimiento académico futuro de los estudiantes. También Orea et al. (2005) realizaron trabajos aplicando técnicas de minería de datos para predecir la deserción escolar en la Universidad Tecnológica de Izúcar de Matamoros (UTIM), tomando como base de análisis los datos del estudio socioeconómico del EXANI-II. Al respecto, Abarca, y Sánchez (2005) analizan la deserción escolar desde la obligatoriedad del empadronamiento a una carrera. Las variables que los autores analizan son la incertidumbre vocacional del estudiante aunado al factor económico. Pautsch et al. (2011) realizaron un proceso de minería de datos con los datos de cohorte desde el 2001 al 2006, aplicados al análisis de la deserción; abordan la temática de la deserción desde el punto de vista académico, social y demográfico. También Icaza et al. (2009) mediante datos recabados en una muestra de estudiantes de nuevo ingreso, utilizando la metodología longitudinal encontraron que existe una correlación baja pero significativa entre el promedio de la preparatoria y el promedio general del primer semestre, así como una diferencia significativa entre las expectativas que tienen los estudiantes sobre su desempeño académico futuro (previo al ingreso) y la realidad y una relación entre el promedio general y la retención.

3. Modelo metodológico CRISP-DM

Se propone la metodología CRISP-DM para modelar la minería de datos, para lo cual se realizó una revisión general de las metodologías, y CRISP-DM fue la más exitosa para realizar los proyectos de investigación. Una de sus principales fortalezas es que sus actividades están ordenadas, ayudando a modelar automáticamente los procesos de la minería de datos, ésta es una subdisciplina de las ciencias de la computación que ha logrado mucho reconocimiento en los últimos años, principalmente porque puede ser usada para diferentes propósitos como herramienta de apoyo en las diferentes disciplinas de las ciencias (Fayyad, 1996). La técnica de minería de datos (MD) puede predecir un porcentaje muy alto de confiabilidad, la probabilidad de desertar de cualquier estudiante.

3.1. Comprensión del Problema

El objetivo del proyecto de minería de datos es descubrir y especificar algunos aspectos importantes de los estudiantes de la Fevaq, referente a las variables que pudieran incidir y afectar su trayectoria académica durante su estancia en la universidad y que resulten como consecuencia en el abandono escolar. Para realizar el análisis se utilizó información académica, institucional, económica y personal de los estudiantes que están en curso en los diferentes semestres.

La información fue procesada y analizada para presentar como producto final una propuesta que ayude a la institución a fortalecer los aspectos que puedan predecir el abandono del estudiante al entrar en la universidad, estos aspectos son considerados como factores de la deserción escolar.

3.2. Evaluación de la situación

Los datos se obtuvieron de la investigación realizada por Pompa (2017) a partir de una encuesta diseñada y aplicada a los alumnos de la Fevaq, para realizar el modelado y el diseño se revisaron y analizaron los datos con el propósito de realizar el proceso de limpieza y preparación de los datos. Para llevar a cabo el modelado de datos del proyecto fue necesario utilizar software especializado como: Weka, IBM Modeler y algunas utilerías de MatLab.

3.3. Determinación de los objetivos de datamining

El objetivo principal del datamining es obtener conocimiento de los datos de la encuesta; con el análisis descriptivo de las variables se pueden determinar ciertas relaciones entre las variables tales como las correlaciones, datos perdidos, datos que son significativamente diferentes a los demás outlier. Estas técnicas sirven para preparar los datos que serán alojados en el modelo a desarrollar.

Al realizar un modelo de minería de datos a través de clasificadores se obtendrán agrupaciones y características comunes de los alumnos de cada semestre, los clasificadores se analizarán y se tomarán decisiones importantes que pudieran ser causas que afecten al alumno en su vida académica.

3.4. Metas de datamining

Se busca encontrar un modelo que sea capaz de predecir las variables que influyen en el abandono de los estudios, a partir de las variables que contiene la encuesta, lo cual implica considerar un modelo bueno aquel que llegue al 80% ($\pm 2\%$) de clasificación, siempre y cuando los clasificados no superen el 5% pues este es el porcentaje de alumnos que finalmente pudieran no abandonar los estudios. El plan para alcanzar los objetivos se muestra en la tabla uno.

Tabla 1 Plan para alcanzar los objetivos de la Minería de Datos

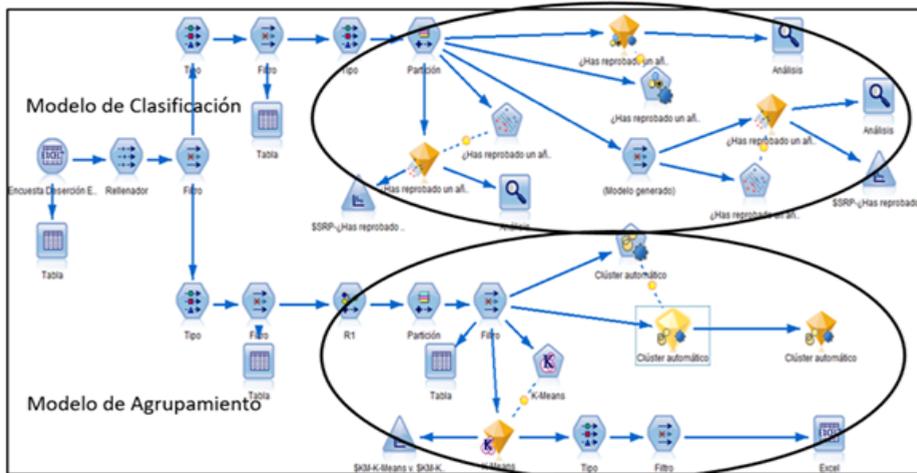
Etapa	Obtención de los datos	Preparación y limpieza de los datos	Validación de los datos	Modelo de minería	Interpretación de resultados
Objetivos	Los datos para el proyecto se obtuvieron del trabajo de Pompa (2017).	Se realizó con varias utilerías, como Clementine, Spss, Matlab. En esta etapa los datos estarán listos para ser encausados al modelo de minería.	Se hizo de acuerdo con el tipo de datos que se traten, para hacer uso de la herramienta correcta.	Son: modelos de clasificación. Estos modelos se llevaron a cabo con el software IBM Modeler.	Realizar la interpretación de los resultados del modelo para obtener el descubrimiento de la información, y obtener un modelo que sea capaz de encontrar los factores que influyen en el abandono del alumno.

Fuente: Elaboración propia

3.5 Modelamiento

El modelo general de MD está construido en dos partes importantes: la primera está conformada por un modelo de clasificación, mientras que en la segunda se analiza a través de un modelo de agrupamiento (Clusters). La primera parte, clasificación, se analizó con el algoritmo SVM (máquinas de soporte vectorial); el modelo de agrupamiento se logró con el algoritmo Kmeas como se muestra en la figura uno.

Figura 1.- Modelo de minería de datos



Fuente: Elaboración propia

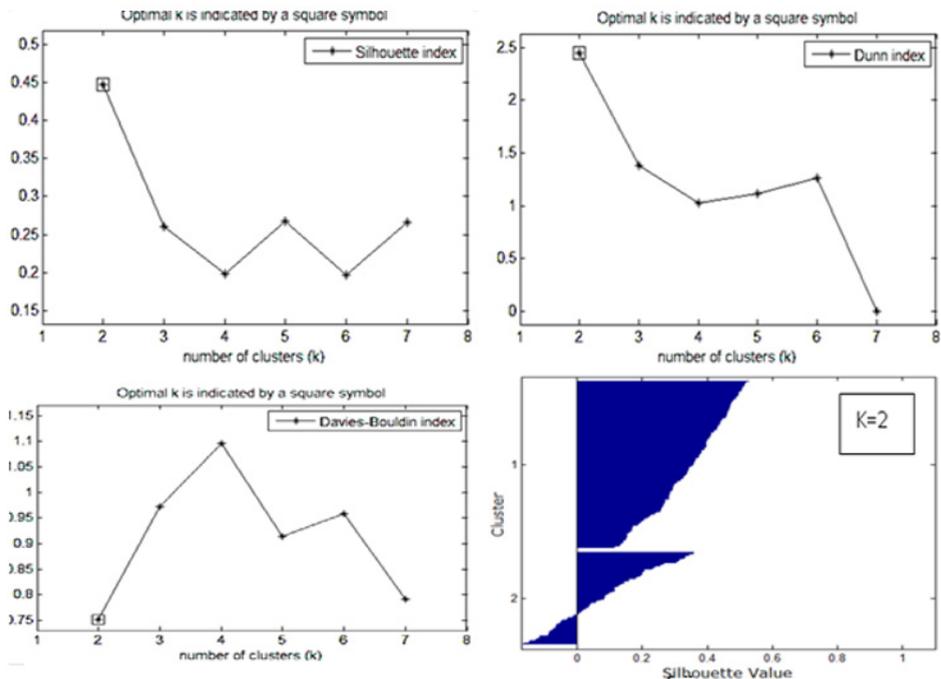
3.6 Evaluación del modelo

El modelo de MD para esta investigación presenta buenos resultados en lo general, la parte del algoritmo de clasificación no presenta ningún problema, sólo en la parte del algoritmo de agrupamiento tiene una desventaja, no se sabe si el número de K 's que proporcionó el nodo automático sea el correcto para los datos, por lo tanto, para saber si son los adecuados se utilizarán las siguientes herramientas de validación.

3.7 Validación de los datos con índices

Para validar los datos se usaron índices de validación que ayudan a evaluar si la formación de los grupos fue la idónea, como se muestra en la figura dos. Los índices fueron, Silueta, Dunn y Davies-Bouldin, los cuales se basan en estimaciones que identifican qué tan compactos o separados se encuentran los grupos formados especialmente por técnicas particionales como la de K-medias (Halkidi et al., 2002).

Figura 2. Validación de los datos con índices



Fuente: Elaboración propia

En la figura dos se pueden apreciar los índices de validación que tienden a optimizar a $k=2$, así también en las gráficas de silueta, se puede ver las siluetas que son más amplias cuando $k=2$. Así, el número de grupos óptimo para este estudio es $k=2$.

En la tabla dos se observa el número óptimo de los grupos sugeridos por los índices de validación, Silueta (valor más alto = 0.44693), Dunn (valor más alto = 2.4472) y Davies-Bouldin (valor más pequeño = 0.75145), para $k=2$. En el caso de silueta con $k=3$ es el valor más próximo a $k=2$ con 0.26711, al existir una diferencia de distancia bastante amplia, se descarta la posibilidad de escoger $k=5$.

Índice de validación	Número de Clúster					
	K= 2	K= 3	K= 4	K= 5	K= 6	K= 7
Silueta	0.44693	0.26015	0.19789	0.26711	0.19715	0.26638
Dunn	2.4472	1.3821	1.0245	1.1135	1.2642	0.00000
Davies-Bouldin	0.75143	0.9725	1.0954	0.91438	0.95871	0.79106

Fuente: Elaboración propia

4. Modelos scoring

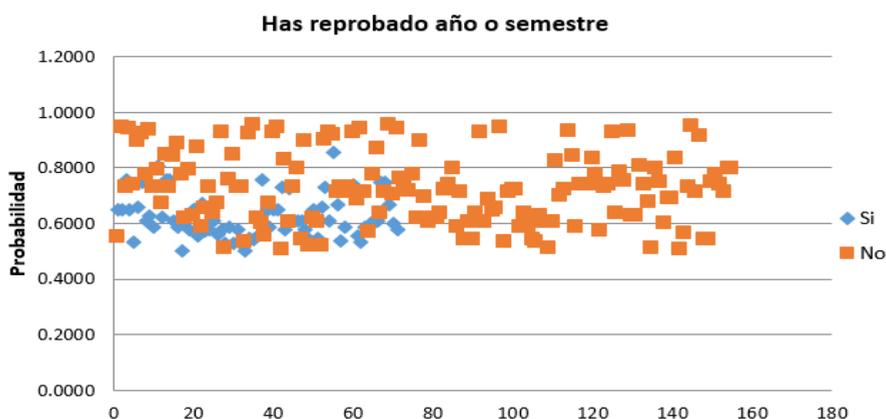
Los modelos de scoring o por puntajes fueron introducidos a partir del año 1970 en el análisis del otorgamiento del crédito, pero generalizados después de 1990 gracias al desarrollo estadístico y tecnológico (Gutiérrez y Alfredo, 2007). Entre los métodos más comunes para la construcción de modelos de scoring se pueden tener en cuenta los modelos LOGIT, PROBIT, Redes Neuronales y el Análisis Discriminante (Ochoa *et al.*, 2010). Los métodos empleados evolucionaron a técnicas matemáticas, econométricas y de inteligencia artificial (Rayo Cantón *et al.*, 2010). El scoring estadístico está basado en historiales de pago, toma los comportamientos pasados para pronosticar comportamientos futuros de los créditos como lo explica Schreiner (2002), un modelo de Scoring utiliza la misma lógica que el analista de crédito, se basa en experiencias y seguimientos de créditos otorgados en el pasado mediante un análisis de las características de los nuevos solicitantes con el fin de calificar o descalificar sus perfiles.

4.1. Modelo de Scoring para determinar la deserción

Para obtener el Scoring es necesario hacer uso de las variables que se requieren para predecir el comportamiento de los alumnos en función de la información proporcionada. Las instituciones educativas deberían hacer uso de esta herramienta para la toma de decisiones mediante un valor crítico, la información recopilada de los alumnos nos permite mediante técnicas estadísticas, árboles de decisión, regresión logística y MVS, saber si el alumno

Los resultados obtenidos del modelo Scoring son las probabilidades de la variable objetivo la cual se considera categórica. En la gráfica uno se observa a los estudiantes que, si han reprobado, representados de color azul; y, los que no han reprobado de color naranja. Podemos apreciar que existen alumnos que han reprobado y tienen una alta probabilidad de no reprobado y viceversa, por lo que es necesario filtrar los datos del campo alumnos que tienen la probabilidad de reprobado.

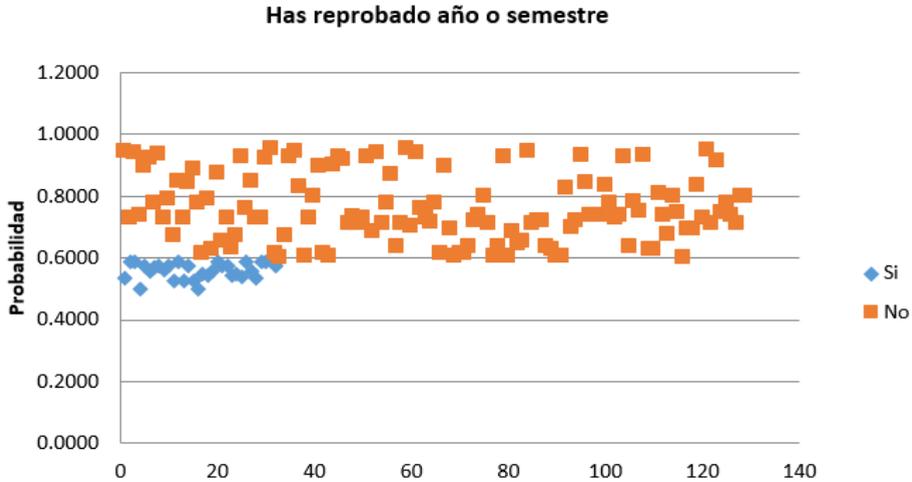
Gráfica 1: Probabilidad reprobado año/semestre



Fuente: Elaboración propia

Una vez realizado el filtro se obtiene solamente el porcentaje de alumnos que tienen la probabilidad de desertar que es el 14.22% del total. También se puede deducir que 24.44% de estudiantes están en el rezago educativo, es decir, los alumnos que han reprobado un año o semestre y persisten en su carrera (gráfica dos).

Gráfica 2: Probabilidad de desertar



Fuente: Elaboración propia

Resultados

Como resultado de este trabajo, el algoritmo de modelo de clasificación, clasificó correctamente el 84% de los datos presentados en la prueba de validación, con la variable dependiente has reprobado un año o semestre en la licenciatura, como se mencionó en el criterio de éxito, al rebasar 80% ($\pm 2\%$); el algoritmo de agrupamiento tiene una media de silueta baja, lo que significa que los grupos no están bien cohesionados, se validaron los datos para las diferentes k 's con los diferentes índices, dando como resultado la mejor validación con $k=2$, la cohesión de los grupos depende de la naturaleza de los datos. El modelo scoring también presenta buenos resultados probabilísticos.

Considerando estos logros en los resultados de desempeño de los modelos de clasificación se puede señalar que las variables más importantes a considerar están relacionadas con el factor económico, esta variable es la más importante en cuestión de la reprobación y deserción en la licenciatura por parte de los alumnos.

Se tienen otras variables poco menos importantes como son las relacionadas al factor académico, también se encontraron otras variables de menor grado como son las relacionadas al factor institucional.

En la comparación de los clusters se pueden apreciar dos grupos que forman parte del Cluster 1 y del Cluster 2. En el primero se obtuvieron mayores puntuaciones y corresponden a las variables más importantes como las de adaptación e identificación con la Fevaq, mientras que en el Cluster 2 se agruparon las menos significativas y que corresponden a las variables académicas.

Por su parte, el predictor sólo muestra las variables de mayor relevancia e interés, pero existen otras variables observables que también son de interés para esta investigación, aunque para el algoritmo son las menos significativas; pero, para esta investigación son muy importantes ya que son las que afectan al sistema educativo y llevan al alumno a reprobar y a desertar. Es posible caracterizar estas variables como las siguientes, ¿Cómo consideras tu nivel de comprensión de lectura?, ¿has asistido a tutorías?, ¿has tenido algún problema de salud?, ¿has tenido asesoría académica fuera del salón de clases por parte de tus profesores?, ¿cómo consideras tu nivel de comprensión de las matemáticas?

Los resultados de los modelos Scoring, permiten obtener la probabilidad de cada alumno a desertar o seguir en la institución, haciendo uso de las variables que se quiera predecir, en ambos modelos se observa el mismo comportamiento. Con las variables predictoras, se obtuvo el puntaje (scoring) asignando a cada estudiante (puntaje o peso), los resultados de acuerdo con las características mencionadas de cada uno quedan de la siguiente manera: el 14.22% están predisuestos a desertar, el 24.44%, se encuentran en el rezago educativo, de acuerdo con el 38.66% de alumnos que contestaron afirmativamente en la encuesta que habían reprobado un año o semestre en la licenciatura. Los resultados del Scoring se asemejan con los datos estadísticos del Sistema Educativo de Michoacán, donde la deserción en el periodo 2012-2014 alcanzó el 10.70% (SNIE, 2014).

Conclusiones

Con los modelos de clasificación utilizados se descubrieron los factores de deserción estudiantil de la Facultad de Economía de la UMSNH. Los factores de deserción más característicos encontrados en el estudio involucran los relacionados, en primer lugar, con el factor económico, representado por la falta de apoyo económico como son las becas, entre otros, por lo cual el alumno se ve obligado a buscar este apoyo en otras organizaciones. Este

aspecto es relevante en la atención del estudiante en la vida académica y las dificultades para lograr con éxito sus estudios, a pesar de que sus padres o tutores los mantienen, la mayoría de los alumnos trabajan simultáneamente a sus estudios. En este sentido se observa que los resultados encontrados en esta parte de la investigación coinciden con los presentados por Del Castillo (2001).

El uso de los modelos Scoring ayudó a encontrar el porcentaje de alumnos que están decididos a retirarse de la institución, característica importante que corresponde al rezago educativo y que es muy habitual en instituciones educativas.

Podemos concluir que las variables de mayor impacto para la deserción o abandono de los estudios de licenciatura en la Fevaq siguen al factor económico, están asociadas al académico, el psicológico y por último la parte institucional.

También podemos señalar que al aplicar el modelo CRISP-DM a una institución educativa puede ayudar al descubrimiento de información para la detección de los factores de deserción, una conclusión importante es que el modelo tiene un proceso de madurez y muestra mejoras en la detección.

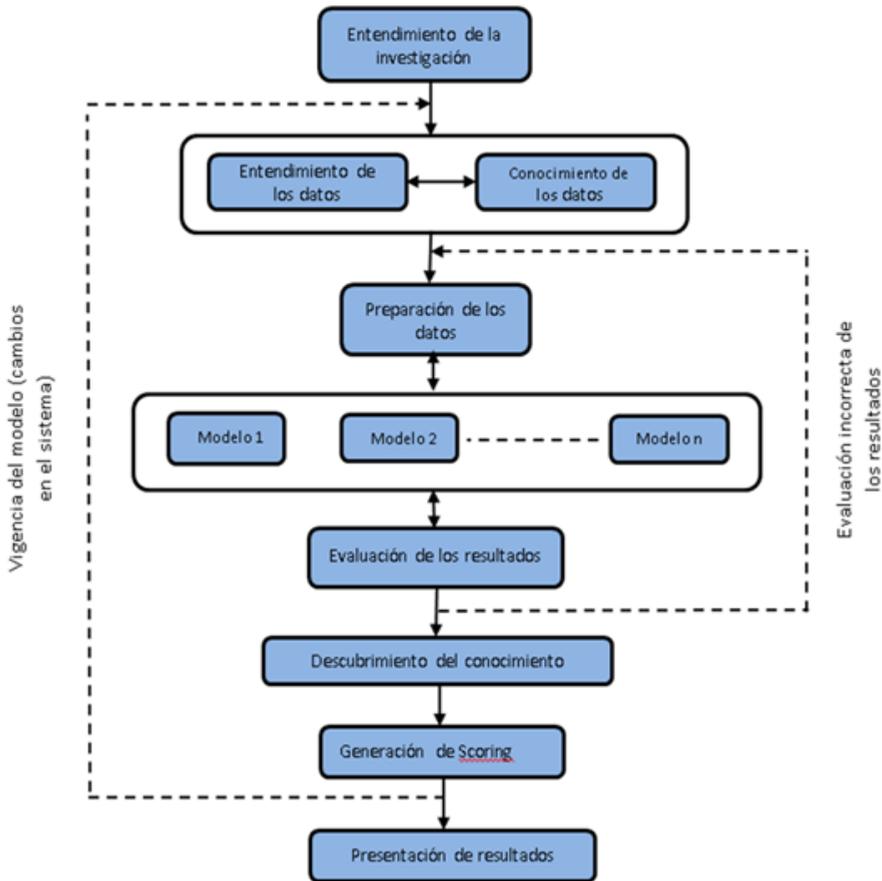
Después de esta revisión general, podrían considerarse las siguientes reflexiones:

- La retención de estudiantes puede generar en términos financieros mayores beneficios económicos para la institución, dado que el estudiante busca una vinculación a largo plazo. En la medida en que el estudiante permanezca en sus estudios de nivel superior hasta culminarlos, el ingreso permanece de acuerdo con lo presupuestado y para ello se requiere inversión principalmente de tres frentes: académico, tecnológico y de servicio personalizado. Planteamiento que coincide con las afirmaciones de Greenberg (2003).
- Diagrama de monitoreo y propuesta de implementación con el objetivo de presentar una propuesta metodológica para detectar las posibles causas de deserción en instituciones de educación superior, las variables obtenidas de los resultados podrán ser monitoreadas cada vez que se requiera a partir de los nuevos datos.

La propuesta se muestra en la figura cuatro, que está orientada a la automatización del sistema y el proceso, con la ayuda de herramientas computa-

cionales, como el desarrollo de una base de datos, warehousing¹, OLTP² y OLAP³, los cuales ayudarán al proceso de transacción, captura, y el análisis de la información.

Figura 4. Modelo metodológico propuesto para la detección de factores de deserción estudiantil



Fuente: Elaboración propia

- 1 El warehousing o almacenamiento es la pieza clave en la gestión de la cadena de suministro, encargándose del almacenamiento, mantenimiento
- 2 OLTP (OnLine Transaction Processing: procesamiento de transacciones en línea) permite el procesamiento de datos rápido y preciso
- 3 El procesamiento analítico en línea (OLAP) es una tecnología que se usa para organizar bases de datos empresariales de gran tamaño y dar soporte a la inteligencia empresarial

BIBLIOGRAFÍA

- Abarca Rodríguez, A. y Sánchez Vindas, M. A. (2005). “La deserción estudiantil en la educación superior: El caso de la Universidad de Costa Rica”. *Redalyc*, 5(4), 1--22.
- Becerril, Guadalupe. (2013) *Deserción escolar, principal reto que enfrenta el sector educativo: SEE y UMSNH - SMRTV*.
<http://smrtv.michoacan.gob.mx/index.php/noticias-sm/desercion-escolar-principal-reto-que-enfrenta-el-sector-educativo-see-y-umsnh>
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J, y Zanasi, A., (1998). *Discovering data mining: from concept to implementation* (Vol. 1): Prentice Hall Upper Saddle River, NJ.
- Catalán, E.. (2011, Septiembre 2013). “Embarazo y escasa habilidad para leer y escribir, causas de deserción escolar: Jara”. *Síntesis de Zacapu*.
<http://www.sintesisdezacapu.com/noticia/articulo,3598/>
- David, Hand, Heikki, Mannila, y Padhraic, Smyth (2001) *Principles of data mining*. MIT Press, MA, 189.
- Del Castillo, A.R. (2001) *Educación Superior y Globalización: Educar ¿para Qué?*. Plaza y Valdés
- Durán, Jorge, y Díaz, Graciela (2001) *Análisis de la deserción estudiantil en la Universidad Autónoma Metropolitana*. En Dobles María y otros autores Investigación en educación. San José Costa Rica: Editorial Universidad Estatal a Distancia-EUNED.
- Espinosa, Héctor Hugo. (2011) “Existe 30% de deserción escolar en la UMSNH” *El Sol de Morelia*. <http://www.oem.com.mx/elsoldemorelia/notas/n2029071.htm>
- Fayyad, Usama, y Uthurusamy, Ramasamy (2002). “Evolving data into mining solutions for insights”. *Communications of the ACM*, 45(8), 28-31.
- Greenberg, Paul (2003). *Las claves de CRM: gestión de relaciones con los clientes*. McGraw-Hill Interamericana de España.
- Halkidi, Maria, Batistakis, Yannis, y Vazirgiannis, Michalis (2002) “Clustering validity checking methods: part II”. *ACM Sigmod Record*, 31(3), 19-27.
- Icaza, Mónica Febles Á, Silva, S. G., y España, A. T. (2009) Deserción en el primer año de la licenciatura: Resultados preliminares. *Consejo Mexicano de Investigación Educativa*. https://www.comie.org.mx/congreso/memoriaelectronica/v10/pdf/area_tematica_16/ponencias/1371-F.pdf

- Martelón, Patricia y Muñiz, E. (1997) “Trayectorias educativas y deserción universitaria en los ochenta”. *México: ANUIES*.
- Oca, Julio Rubio (2007) “La evaluación y acreditación de la educación superior en México: Un largo camino aún por recorrer”. *Reencuentro* (050), 35--44.
- Ochoa, P, Camilo, Juan, Galeano, M, Agudelo, V, y Gabriel, Luis (2010) “Construcción de un modelo de scoring para el otorgamiento de crédito en una entidad financiera”. *Perfil de Coyuntura Económica* (16), 191-222.
- Orea, Sergio V., Vargas, A. S., y Alonso, M. G. (2005) “Minería de datos: Predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos”. *Ene*, 779(73), 33.
- Pautsch, J., Kuna, Horacio, y Godoy, Antonia E. (2011) “Resultados preliminares del proceso de minería de datos aplicado al análisis de la deserción en carreras de informática utilizando herramientas open source”. En XVII Congreso Argentino de Ciencias de la Computación.
- Pompa, Y. (2017). *Detección de las Posibles Causas de Deserción Escolar en la Facultad de Economía "Vasco de Quiroga" de la UMSNH*. (Tesis de Maestría en Tecnologías de la Información), Universidad de Morelia.
- Porcel, Eduardo, Dapozo, Gladys N, y López, María Victoria. (2009) “Modelos predictivos y técnicas de minería de datos para la identificación de factores asociados al rendimiento académico de alumnos universitarios”. Paper presented at the XI Workshop de Investigadores en Ciencias de la Computación.
- Quadratin, Agencia de noticias. (2024) “Deserción en UMSNH, de 30%; mayoría, de ingenierías y ciencias exactas”. *Quadratin*. <https://www.quadratin.com.mx/educativas/desercion-en-umsnh-de-30-mayoria-de-ingenierias-y-ciencias-exactas/>
- Rayo Cantón, Salvador, Lara Rubio, Juan, y Camino Blasco, David (2010) “Un modelo de Credit Scoring para instituciones de microfinanzas en el marco de Basilea II”. *Journal of Economics, Finance and Administrative Science*, 15(28), 89-124.
- Rodríguez, M. G. O. (2018) “Deserción de estudiantes de licenciatura de la UMSNH Análisis y propuesta de solución”. *Economía y Sociedad*, 22(38), 15-32.
- Schreiner, Mark (2002) “Ventajas y desventajas del scoring estadístico para

- las microfinanzas”. *Microfinancierisk Management*, 1-40.
- SNIE. (2014). “Estadística del Sistema Educativo Michoacán Ciclo Escolar 2013-2014”. *Estadística e indicadores educativos*. http://www.sniesep.gob.mx/descargas/estadistica_e_indicadores/estadistica_e_indicadores_educativos_16MICH.pdf
- Tinto, Vicent (1989) “Definir la deserción: Una cuestión de perspectiva”. *Revista de Educación Superior*, 71, pp. 33--51.
- Tünnermann Bernheim, C. (1996) *La Educación Superior en el Umbral del Siglo XXI*. CRESALC/UNESCO.
- Velazquez, y Romero, Lázaro (2012) “Deserción escolar en las IES públicas”. Recuperado el 3 de marzo de 2012 de <http://www.peu.buap.mx/Revista21/Articulos/desercion%20de%20ies.pdf>
- Villavicencio, Záyin Dáleth (2013) “Preocupante, la deserción escolar en Michoacán, alertan SEE y UMSNH: Sólo la mitad de niños que ingresan a primaria terminará la secundaria, detallan”. *Periódico La Jornada*. <http://www.lajornadamichoacan.com.mx/2013/05/17/solo-la-mitad-de-ninos-que-ingresan-a-primaria-concluye-la-secundaria/>

ACERCA DE LOS AUTORES

Arturo García Campos es Ingeniero Eléctrico por la UMSNH, Maestro en Tecnologías de la Información por la Universidad de Morelia (UdeM). Cuenta con 25 años de experiencia docente en la Universidad Michoacana de San Nicolás de Hidalgo (UMSNH), en la Facultad de Economía, impartiendo las asignaturas de computación aplicada, matemáticas y estadística. Además, ha contribuido con publicaciones en revistas académicas.

Yiringari Pompa Flores es Licenciada en Informática por la Universidad de Morelia (UDEM), Maestra en Tecnologías de Información por la misma universidad. Trabajó durante cinco años en el área de laboratorios de cómputo y soporte en la Universidad Latina de América. Cuenta con 17 años de experiencia en la Universidad Michoacana desempeñándose en el área de cómputo de la Facultad de Economía impartiendo la materia de Computación Aplicada.