



EXPERIENCIAS EN EL DESARROLLO DE UN PROYECTO PARA EL ESTUDIO DE LA DESERCIÓN ESCOLAR CON HERRAMIENTAS DE MINERÍA DE DATOS

Manuel Sabino Lazo Cortés

*Tecnológico Nacional de México/IT Tlalnepantla
manuel.lc@tlalnepantla.tecnm.mx*

Abstract

Some experiences are presented based on the analysis of the student database with the objective of studying school dropout. Aspects related to the subject are described, some characteristics of the information are discussed and some conclusions are presented. No manifestations were observed that allow us to affirm that school dropout occurs more in female students than in male students

Palabras clave: deserción escolar; minería de datos; patrones frecuentes

El presente proyecto tiene como objetivo realizar un análisis sobre el fenómeno de la deserción escolar dentro del Instituto Tecnológico de Tlalnepantla a través de un estudio para la identificación de posibles causas del abandono escolar.

La deserción escolar es un fenómeno que afecta a todos los niveles educativos. Varios autores han afirmado que el problema del abandono de los estudios debe examinarse en el contexto social y económico de la región, en particular en el contexto del sistema educativo en su conjunto. Esto se debe a la importancia de los factores sociales que influyen en los individuos, provocando que tomen decisiones educativas. Es importante hacer énfasis en el

hecho que en la escuela no solamente se llevan a cabo los diversos procesos educativos, sino que a su vez permite crecer en valores, aprender a socializar, desarrollar habilidades cognitivas y físicas etc.

Referente a los alumnos, es sumamente notorio que en un alto porcentaje de casos los alumnos que conllevan riesgos de deserción cuentan con un perfil en específico en relación con su aprovechamiento escolar, esto de la mano con la relatividad de la asistencia, las calificaciones obtenidas en sus periodos parciales lo cual se engloba bajo el concepto de desaprovechamiento escolar.



Para tener una estimación exacta de las causas del abandono escolar, se deberán realizar los análisis pertinentes y detallados de los datos sobre los perfiles de la comunidad estudiantil para obtener la incidencia mayor en la deserción incluyendo el análisis respecto al género.

El abandono escolar es sin duda una preocupación para las autoridades gubernamentales y autoridades educativas; por ello, este proyecto se desarrolla para el aporte a la institución analizada.

La mayoría de investigaciones relacionadas en este tema son meramente cuantitativas con el fin de resultar las variables y factores del fenómeno, dejando a un lado las circunstancias causales del problema.

Las consecuencias de la deserción escolar conllevan a distintas afectaciones en todo el contexto social, por ejemplo:

Incidir negativamente en el desarrollo del capital humano de un país.

Aumentar los costos sociales porque el capital humano está menos cualificado.

Obtención de un empleo que proporciona un salario bajo o insuficiente para cubrir las necesidades básicas.

Limitar el desarrollo económico y social de los países.

Fomentar la desigualdad social. Los índices de actividad delictiva tienden a aumentar.

Aumenta la posibilidad de realizar tareas incorrectamente.

Al no haber completado los estudios o la carrera, se puede experimentar depresión y frustración personal.

Aumentar el número de personas que dependen de las prestaciones del gobierno.

Un informe de 2013 señala que México ocupó el primer lugar en el número de adolescentes de 15 a 18 años que abandonan los estudios, el grupo de edad con más probabilidades de terminar un bachillerato y asistir a la universidad. Y ratificó el tercer lugar entre las naciones con mayor población juvenil que no estudia ni trabaja, con 7 millones 337 mil 520, condiciones calificadas como un drama y algo cruel por una falta estructural por la Organización para la Cooperación y el Desarrollo Económico (OCDE). En relación con lo anterior, cabe destacar que la cuestión del abandono escolar está presente en diversos contextos desde hace muchos años.

Muchos estudiantes se estresan porque tienen obligaciones externas que les impiden centrar el 100% de su atención en los estudios. Como consecuencia, buscan el camino más sencillo, que incluye la deserción escolar y el abandono de sus estudios.

Las instituciones de educación superior cuentan con sistemas de información académica que registran, entre otros, datos personales, socioeconómicos y de rendimiento académico de los estudiantes antes y durante su estancia en la institución, como es el caso del Instituto Tecnológico de Tlalnepantla.

Para contribuir con el conocimiento de los factores que influyen al problema de la deserción estudiantil se pretende apoyarse de la herramienta de minería propuesta "STreeDCMiner" [9,10], el cual permite procesar datos de diferente naturaleza, sin necesidad de usar códigos, o de discretizar los datos previamente, a la vez que presenta una amplia variedad de criterios de comparación



entre los valores de los diferentes atributos a considerar, para luego evaluar los resultados obtenidos en la investigación.

En este trabajo se expone la experiencia de cuatro estudiantes de Ingeniería en Tecnologías de la Información y Comunicaciones durante su residencia profesional, utilizando el minero STreeDCMiner.

base de datos de estudiantes del SII (Sistema Integral de Información). A continuación, la Tabla 1 presenta una breve descripción de las tablas que inicialmente se consideraron. La llave primaria en todas ellas es el número de control del estudiante. Todo el trabajo se realizó sin acceso a la información primaria de identificación del estudiante (nombre y apellidos).

Desarrollo

Para el desarrollo de este trabajo se consideró la información almacenada en la

Tabla 1. Relación de tablas con información del estudiante consideradas inicialmente, con sus campos

alumnos	alumnos_generales	alumnos_socioeconomicos	bajas_materias	calificaciones parciales	historia_alumno	Selecciones_materias
<i>no. de control</i>						
carrera	entidad_federativa	con_quien_vives	periodo	periodo	periodo	periodo
reticula	código_postal	nivel_estudios_padre	grupo	grupo	grupo	grupo
especialidad	hijos	nivel_estudios-madre	materia	materia	materia	materia
nivel_escolar	discapacidad	de_quien_dependes	carrera	unidad	calificacion	calificacion
semestre	dependientes		estatus	calificacion	tipo_evaluacion	Tipo-evaluacion
estatus					period_acredita	repeticion
plan_de_estudios						no_presento
sexo						
periodo_ingreso						

La Tabla 2 contiene las carreras que se consideraron en el presente estudio, así como sus diferentes retículas. Las retículas se diferencian entre otras cosas por el número

total de créditos. En algunas carreras no se dispuso de información para todas las retículas, lo cual puede apreciarse en las tablas subsiguientes.



Tabla 2. Relación de carreras (y retículas) incluidas en la información de la que se dispuso

CARRERA	RETÍCULA	CRÉDITOS
LICENCIATURA EN ADMINISTRACIÓN	1	440
LICENCIATURA EN ADMINISTRACIÓN	2	392
INGENIERÍA ELÉCTRICA	1	440
INGENIERÍA ELÉCTRICA	2	400
INGENIERÍA ELÉCTRICA	3	260
INGENIERÍA ELECTROMECAÁNICA	1	440
INGENIERÍA ELECTROMECAÁNICA	2	400
INGENIERÍA ELECTROMECAÁNICA	3	260
INGENIERÍA INDUSTRIAL	1	430
INGENIERÍA INDUSTRIAL	2	400
INGENIERÍA INDUSTRIAL	3	260
INGENIERÍA MECATRÓNICA	1	400
INGENIERÍA MECATRÓNICA	2	260
LICENCIATURA EN INFORMÁTICA	1	400
LICENCIATURA EN INFORMÁTICA	2	438
INGENIERÍA MECÁNICA	1	440
INGENIERÍA MECÁNICA	2	400
INGENIERÍA MECÁNICA	3	260

ING. EN TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIONES	1	260
ING. EN TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIONES	2	260
INGENIERÍA EN ADMINISTRACIÓN	1	260
INGENIERÍA EN GESTIÓN EMPRESARIAL	1	260
INGENIERÍA EN GESTIÓN EMPRESARIAL	2	260

La Tabla 3 contiene los valores que se consideraron para los campos 'con quien vives', 'de quien dependes' y 'estudios padres'.

Tabla 3. Valores para los campos 'con quien vives', 'de quien dependes' y 'estudios padres'

con quien vive	de quien depende	estudios padres
Padre y Madre	Padre y Madre	No lo se
Padre	Padre, Madre y Yo mismo	No sabe leer ni escribir
Madre	Padre	No fue a la escuela
Hermanos	Padre y Yo mismo	No terminó la Primaria
Cónyuge o pareja	Madre	Terminó la Primaria
Otro familiar	Madre y Yo mismo	Tiene alguna capacitación técnica después de la Primaria
Amigo o amigos	Hermanos	No terminó la Secundaria



Sólo	Hermanos y Yo mismo	Terminó la Secundaria
Hijos	Cónyuge o pareja	Tiene alguna capacitación técnica después de la Secundaria
Otro	Cónyuge, pareja y Yo mismo	Tiene estudios de técnico profesional incompletos
	Otro familiar o amigo	Tiene estudios de técnico profesional completos
	Yo mismo	No terminó la Preparatoria o Bachillerato
	Otro	Terminó la Preparatoria o Bachillerato
		No terminó la Licenciatura, Ingeniería o Normal
		Terminó la Licenciatura Ingeniería o Normal
		No terminó la Maestría o Doctorado
		Terminó la Maestría o Doctorado
		Otro

La Tabla 4 contiene los valores del campo status, los cuales aparecen clasificados según se considere que el estudiante desertó o no de los estudios. Esta clasificación fue elaborada y adoptada a partir de la opinión de los participantes en el proyecto.

Tabla 4. Valores del campo 'status' agrupados en deserción o no deserción

STATUS	DESERCION
Activo	no
Inscrito	no
Egresado	no
Titulado	no
Baja temporal por Reglamento	no
Baja definitiva por Traslado	no
Baja temporal por Movilidad	no
Baja temporal Voluntaria	si
Baja temporal por adeudo de exámenes especiales	si
Baja temporal por no inscripción	si
Baja definitiva voluntaria	si
Baja definitiva por reprobación en primer semestre	si
Baja definitiva por examen especial reprobado	si
Baja definitiva por alcanzar el número máximo semestres permitido	si
Baja definitiva por no inscripción a segundo semestre	si
Baja definitiva por no cumplir con requisito de inscripción	si
Baja definitiva por reprobación de curso especial	si
Baja definitiva por exceder cursos especiales	si
Baja Definitiva por Inactividad en el Sistema	si

La Tabla 5 contiene el listado de especialidades por carreras y retículas.



Tabla 5. Listado de especialidades por carreras y retículas

CARRERA	RETÍCULA	ESPECIALIDAD
Ingeniería Eléctrica	1	Aplicaciones Industriales
Ingeniería Eléctrica	2	Automatización Industrial y de Servicios
Ingeniería Eléctrica	3	Generación y Operación de Sistemas Eléctricos
Ingeniería Eléctrica	3	Ingeniería Eléctrica
Ingeniería Eléctrica	3	Sistemas Eléctricos de Potencia I. R. y E.
Ingeniería Electromecánica	1	Manufactura
Ingeniería Electromecánica	2	Automatización
Ingeniería Electromecánica	3	Mantenimiento Industrial
Ingeniería Electromecánica	3	Transformación de Procesos Productivos
Ingeniería en Administración	1	Dirección de Procesos Sustentables
Ingeniería en Administración	1	Gerencia de Sostenibilidad
Ingeniería en Gestión Empresarial	1	Habilidades Directivas
Ingeniería en Gestión Empresarial	1	Dirección Comercial
Ingeniería en Gestión Empresarial	2	Habilidades Directivas
Ingeniería en Gestión Empresarial	2	Dirección Comercial
Ingeniería en Tecnologías de la Información y Comunicaciones	1	Creación e implementación de sw con enfoque dual
Ingeniería en Tecnologías de la Información y Comunicaciones	1	Ingeniería de Software
Ingeniería en Tecnologías de la Información y Comunicaciones	1	Gobernabilidad en TIC's
Ingeniería en Tecnologías de la Información y Comunicaciones	1	Gestión y Gobierno
Ingeniería en Tecnologías de la Información y Comunicaciones	1	Aplicaciones Web
Ingeniería en Tecnologías de la Información y Comunicaciones	2	Ingeniería de Software
Ingeniería en Tecnologías de la Información y Comunicaciones	2	Gobernabilidad en TIC's
Ingeniería en Tecnologías de la Información y Comunicaciones	2	Gestión y Gobierno
Ingeniería en Tecnologías de la Información y Comunicaciones	2	Aplicaciones Web
Ingeniería Industrial	1	Calidad y Productividad
Ingeniería Industrial	2	Sistemas de Producción Calidad y Logística
Ingeniería Industrial	2	Artes Gráficas
Ingeniería Industrial	3	Automatización Sustentabilidad E Innovación



Ingeniería Industrial	3	Sistema de Producción Sustentable e Innovación
Ingeniería Industrial	3	Sistemas de Logística
Ingeniería Mecánica	1	Manufactura
Ingeniería Mecánica	2	Diseño y Manufactura
Ingeniería Mecánica	3	Diseño y Manufactura
Ingeniería Mecánica	3	Sistemas Integrados de Diseño y Manufactura
Ingeniería Mecatrónica	1	Integración de Sistemas de Manufactura Industrial
Ingeniería Mecatrónica	2	Automatización de Sistemas y Producción
Ingeniería Mecatrónica	2	Proyectos de Ingeniería Mecatrónica
Licenciatura en Administración	2	Administración de Negocios Nacionales e Internacionales
Licenciatura en Informática	1	Sistemas Distribuidos con Optativa I
Licenciatura en Informática	1	Sistemas Distribuidos Optativa II
Licenciatura en Informática	2	Redes
Licenciatura en Informática	2	Sistemas Administrativos

La población estudiantil de la que se tenía información ascendió a 21801 estudiantes, agrupados por carreras, según la Tabla 6, además aparecen desglosados según el género.

La agrupación por los perfiles se hizo para dividir el proyecto entre los estudiantes participando de la residencia profesional y balancear la carga de trabajo entre ellos.

Tabla 6. Estudiantes con datos por carrera

carrera	total	masculino	femenino	perfil
ING. INDUSTRIAL	4422	3223	1199	INDUSTRIAL
ING. EN ADMINISTRACIÓN	769	384	385	EMPRESARIAL
ING. EN GESTIÓN EMPRESARIAL	4280	1938	2342	
LIC. EN ADMINISTRACIÓN	1570	593	977	
ING. ELÉCTRICA	1027	955	72	ELECTRO MECÁNICO
ING. ELECTROMECAÁNICA	1129	1058	71	
ING. MECÁNICA	1458	1401	57	
ING. MECATRÓNICA	2457	2188	269	
LIC. EN INFORMÁTICA	1529	916	613	INFORMÁTICO
ING. EN TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIONES	3160	2278	882	



En las siguientes tablas (Tabla 7 a-q) se muestra la composición por género en los estudiantes por carrera y retícula, considerando el grupo completo, el de desertores y el de no desertores. La tabla se refiere a las carreras y retículas de las cuales se tuvo información.

Tabla 7 a Composición por género (Mecatrónica 2)

CARRERA						RETÍCULA
INGENIERÍA MECATRÓNICA						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
1657	214	486	52	1171	162	
%	11.44	%	9.67	%	12.15	

Tabla 7 b Composición por género (Mecatrónica 1)

CARRERA						RETÍCULA
INGENIERÍA MECATRÓNICA						1
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
531	55	263	19	268	36	
%	9.39	%	6.74	%	11.84	

Tabla 7 c Composición por género (Industrial 3)

CARRERA						RETÍCULA
INGENIERÍA INDUSTRIAL						3
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
2239	856	751	170	1488	686	
%	27.66	%	18.46	%	31.55	

Tabla 7 d Composición por género (Industrial 2)

CARRERA						RETÍCULA
INGENIERÍA INDUSTRIAL						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
984	343	513	166	471	177	

%	25.85	%	24.45	%	27.31
---	-------	---	-------	---	-------

Tabla 7 e Composición por género (Electromecánica 3)

CARRERA						RETÍCULA
INGENIERÍA ELECTROMECAÁNICA						3
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
730	62	309	23	421	39	
%	7.83	%	6.93	%	8.48	

Tabla 7 f Composición por género (Electromecánica 2)

CARRERA						RETÍCULA
INGENIERÍA ELECTROMECAÁNICA						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
328	9	172	5	156	4	
%	2.67	%	2.82	%	2.50	

Tabla 7 g Composición por género (Eléctrica 3)

CARRERA						RETÍCULA
INGENIERÍA ELÉCTRICA						3
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
678	55	266	14	412	41	
%	7.50	%	5.00	%	9.05	

Tabla 7 h Composición por género (Eléctrica 2)

CARRERA						RETÍCULA
INGENIERÍA ELÉCTRICA						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
277	17	155	11	122	6	
%	5.78	%	6.63	%	4.69	



Tabla 7 i Composición por género (Mecánica 3)

CARRERA						RETÍCULA
INGENIERÍA MECÁNICA						3
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
1038	47	252	12	786	35	
%	4.33	%	4.55	%	4.26	

Tabla 7 j Composición por género (Mecánica 2)

CARRERA						RETÍCULA
INGENIERÍA MECÁNICA						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
363	10	172	7	191	3	
%	2.68	%	3.91	%	1.55	

Tabla 7 k Composición por género (Informática 1)

CARRERA						RETÍCULA
LICENCIATURA EN INFORMÁTICA						1
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
916	613	494	279	422	334	
%	40.09	%	36.09	%	44.18	

Tabla 7 l Composición por género (Lic. Administración 2)

CARRERA						RETÍCULA
LICENCIATURA EN ADMINISTRACIÓN						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
593	977	290	371	303	606	
%	62.23	%	56.13	%	66.67	

Tabla 7 m Composición por género (IGE 2)

CARRERA						RETÍCULA

INGENIERÍA EMPRESARIAL EN GESTIÓN						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
207	270	41	22	166	248	
%	56.60	%	34.92	%	59.90	

Tabla 7 n Composición por género (IGE 1)

CARRERA						RETÍCULA
INGENIERÍA EMPRESARIAL EN GESTIÓN						1
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
1731	2072	706	641	1025	1431	
%	54.48	%	47.59	%	58.27	

Tabla 7 o Composición por género (Ing. Administración 1)

CARRERA						RETÍCULA
INGENIERÍA EN ADMINISTRACIÓN						1
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
384	385	75	59	309	326	
%	50.07	%	44.03	%	51.34	

Tabla 7 p Composición por género (ITICs 2)

CARRERA						RETÍCULA
INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIONES						2
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
316	133	92	27	224	106	
%	29.62	%	22.69	%	32.12	

Tabla 7 q Composición por género (ITICs 1)

CARRERA						RETÍCULA



INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIONES						1
TODOS		DESERTORES		NO DESERTORES		
M	F	M	F	M	F	
1962	749	798	287	1164	462	
%	27.63	%	26.45	%	28.41	

A partir de las Tablas 7 a-q puede apreciarse que solamente en tres de ellas el porcentaje de mujeres que desertan es mayor que el porcentaje de mujeres en el conjunto total de la carrera y retícula correspondiente; pero en ningún caso la diferencia entre esos porcentajes rebasa el 2%, por lo que puede afirmarse a partir de los datos disponibles que la deserción escolar no es un fenómeno que se manifieste de una manera más frecuentemente entre las estudiantes que entre los estudiantes.

En términos absolutos se presenta una cifra de índice de deserción del 36.69 si se consideran de manera general todas las carreras y retículas; con un mínimo de 13.2 y un máximo de 56.46.

A continuación, se presenta información acerca de la calidad de la información.

Para el tema de estudio, los datos socioeconómicos del estudiante parecían ser de gran interés, no obstante, hubo que prescindir de ellos al comprobar que solamente el 4.5% de los estudiantes poseían información relativa a los campos “con_quien_vives”, “de_quien_dependes”, “nivel-estudios-padre” y “nivel-estudios_madre”. Por otro lado, el campo “dependientes” no contenía información para ningún estudiante.

La tabla “historia_alumno” contiene la información de la trayectoria del estudiante con cada una de las calificaciones obtenidas en cada materia en cada oportunidad que fue cursada. El

estudio de esta tabla nos permitió obtener información acerca de aquellas materias que pudieran estar más relacionadas con el fracaso estudiantil. El procesamiento se realizó con el minero STreeDCMiner, ya que en el mismo resulta fácil poder procesar datos de diferente naturaleza, ya sean numéricos o categóricos y de una manera fácil modificar los parámetros respecto a los patrones frecuentes y las reglas de asociación. Aunque esta parte constituyó el mayor volumen de trabajo, los resultados obtenidos requieren de otros análisis y serán objeto de un trabajo futuro.

Con el desarrollo de este trabajo se pretendía además evaluar la pertinencia del uso de STreeDCMiner en tareas como esta. Simultáneamente se pudieron desarrollar algunas recomendaciones para un uso más eficiente del minero. Las recomendaciones fundamentales son:

1. Proveer al software minero STreeDCMiner de una interfaz, que sea amigable para el usuario, debido a que no cuenta con una interfaz que permita hacer uso del software de una forma más ágil y eficiente. Esta interfaz puede contar con accesos que permitan hacer la carga directa a una base de datos, para evitar la pérdida de datos en el momento de la creación de los formatos de entrada específicos. Aunado a esto la interfaz también puede contar con diferentes accesos que permitan desplazarte en más ventanas dentro del minero, esto con la finalidad de que se puedan observar gráficos que ilustren los resultados de los patrones frecuentes, así como las reglas de asociación, además de que dentro del minero se pueda realizar directamente la creación de los formatos específicos, con la finalidad de que se pueda realizar una mejor gestión del tiempo, además de un



aporte para evitar hacer el uso de otros programas externos para la creación de estos.

2. Una segunda recomendación es relativa a los mensajes que arroja el software minero STreeDCMiner a los usuarios, ya que los mensajes que el minero mostró a lo largo del proyecto fueron un poco confusos, por lo cual si a esto se le sumara la inclusión de un manual de usuario, que pueda apoyarlo en la interpretación de cada uno de los mensajes de error que pueda presentar el minero a lo largo del uso del mismo, sería una aportación gratificante para el usuario que apenas se vaya familiarizando con el uso del minero.
3. La tercera recomendación para el software minero es la evaluación del tiempo empleado para la ejecución de los procesos de minería, ya que los procesos que se realizaron a lo largo del estudio, se comprobaron que al manejar diferentes parámetros de entrada con abundantes datos, requerían procesos extremadamente demorados, para que al final del proceso, el software minero, arrojara un resultado negativo, de limitaciones de proceso (por falta de memoria, por ejemplo) ya que estos influían en el tiempo de respuesta de la obtención de los formatos de salida.
4. Informar de los límites de procesamiento de datos, (el espacio requerido de memoria en JAVA) para ejecutar el proceso, así como el número de información contenida en los formatos específicos de entrada.
5. Aunado a las recomendaciones anteriores, la inclusión de seguridad dentro del software, debido a que como es un programa que trabaja con grandes volúmenes de información estos pueden ser blanco de distintos ataques como: DoS (denial-of-service), R2L, U2R, Probing.

Conclusiones

Las instituciones de educación constituyen células importantes en la construcción del tejido social. Su labor va más allá de la formación de profesionales; y por ello el abandono escolar no puede ser visto únicamente a través de su incidencia en el número de profesionales formados, sino que tiene importantes implicaciones en el desarrollo socio-cultural de un país. En este estudio se constató que en algunas carreras los índices de deserción son bastante altos, llegando al 50.6%. No fue posible profundizar en posibles causas de carácter socioeconómico debido a la mala calidad de la información, ya que en algunos casos menos del 5% de los estudiantes poseían información. Se pudo concluir que en los datos analizados no había una diferenciación que no permitiera afirmar que el abandono de estudio se presentaba más en el género femenino que en el masculino, lo cual contradice algunos estudios publicados respecto a este fenómeno en otros ámbitos. Se obtuvo información específica respecto a la correlación que podría haber entre algunas materias y el abandono de estudio, pero dichos resultados no son concluyentes y requieren de nuevas fases y nuevos tipos de estudios de carácter más multifactorial. El uso del minero STreeDCMiner permitió elaborar un conjunto de recomendaciones para mejorar su desempeño y facilitar su uso. Este trabajo se desarrolló como Residencia Profesional y contribuyó al desarrollo de habilidades y competencias en los futuros profesionales.

Referencias

- [1]. Alania Ricaldi, P. F. (2018). (Tesis) Aplicación de técnicas de minería de datos para predecir la deserción estudiantil de la Facultad de Ingeniería



- de la Universidad Nacional Daniel Alcides Carrión. Perú.
- [2]. Amaya Torrado, Y. K., Barrientos Avendaño, E. y Heredia Vizcaíno, D. J. (2014). Modelo predictivo de deserción estudiantil utilizando técnicas de minería de datos. <https://documentos.redclara.net/bitstream/10786/759/1/124-22-3-2014-Modelo%20predictivo%20de%20deserción%20estudiantil%20utilizando%20técnicas%20de%20minería%20de%20datos.pdf>. Colombia.
- [3]. Brito Guzmán, M. P. (2012). Modelos predictivos y explicativos del rendimiento académico universitario: caso de una institución privada en México. E-Prints Complutense Repositorio Institucional de la UCM: <https://eprints.ucm.es/id/eprint/15335/>
- [4]. Dzay Chulim, F. y Narváez Trejo, O. M. (2012). La deserción escolar desde la perspectiva estudiantil Universidad Veracruzana.
- [5]. Lladó Lárraga, D. M. y Mares Rodríguez, H. A. (2017). Factores que impactan la deserción escolar: percepción de los estudiantes de la escuela preparatoria federalizada no.1 ing. Marte R. Gómez. Congreso Nacional de Investigación Educativa (CNIE). San Luis Potosí. 2017.
- [6]. López Ávila, A. L. (2020) Deserción escolar de alumnos de la carrera de licenciado en geografía UAEMEX, plan 2017: Una mirada desde la perspectiva geográfica. Tesis. Univ Aut. del edo. de México (UAEM).
- [7]. López Villafaña, L. y Beltrán Solache, A. (2018). Causas de deserción en estudiantes de educación superior: el caso de la licenciatura de contaduría. *Pistas Educativas*, 33(102) 24-39.
- [8]. Piscocoya Ordoñez, L. E. (2016). Aplicación de técnicas de minería de datos para predecir la deserción estudiantil en la educación básica regular en la región de Lambayeque. Chiclayo, Perú. (Tesis) <https://repositorio.uss.edu.pe/bitstream/handle/20.500.12802/4066/Tesisterminada.pdf?sequence=1&isAllowed=y>
- [9]. Rodríguez-González, A. Y., Martínez-Trinidad, J. F., Carrasco-Ochoa, J. A. and Ruiz-Shulcloper, J. (2011). RP-Miner: a relaxed prune algorithm for frequent similar pattern mining. *Knowledge and information systems*, 27(3), 451-471.
- [10]. Rodríguez-González, A. Y., Martínez-Trinidad, J. F., Carrasco-Ochoa, J. A. and Ruiz-Shulcloper, J. (2013). Mining frequent patterns and association rules using similarities. *Expert Systems with Applications*, 40(17), 6823-6836.
- [11]. Rodríguez Lagunas, J. y Leyva Piña, M. A. (2007). La deserción escolar universitaria. La experiencia de la UAM. Entre el déficit de la oferta educativa superior y las dificultades de la retención escolar. *El cotidiano*, 22(142), 98-111.