

UNIVERSIDAD
PRIVADA
DEL NORTE

FACULTAD DE INGENIERÍA

Carrera de Ingeniería Industrial

“ANÁLISIS DE MODELOS PREDICTIVOS QUE ESTIMEN LA DESERCIÓN DE ESTUDIANTES DE PREGRADO”: una revisión de la literatura científica.

Trabajo de investigación para optar al grado de:

Bachiller en Ingeniería Industrial

Autores:

César Edilberto Rivas Guerra

Asesor:

Mg. NINAQUISPE SOTO, MARIO EDISON

Lima - Perú

2019



DEDICATORIA

El presente trabajo está dedicado para mis familiares, que desde un primer momento me brindaron su apoyo incondicional para lograr este gran reto, muchas gracias a ustedes.



AGRADECIMIENTO

El agradecimiento a todos los docentes que acompañaron en todo este proceso de aprendizaje, mucho de lo aprendido en aulas serán expuestas en el presente trabajo.



Tabla de contenido

DEDICATORIA.....	2
AGRADECIMIENTO	3
ÍNDICE DE TABLAS	5
ÍNDICE DE FIGURAS.....	6
RESUMEN.....	7
CAPÍTULO I. INTRODUCCIÓN	8
CAPÍTULO II. METODOLOGÍA.....	10
CAPÍTULO III. RESULTADOS	23
CAPÍTULO IV. CONCLUSIONES	28
REFERENCIAS	29



ÍNDICE DE TABLAS

Tabla 1. Base de datos de revisiones de literaturas científicas parte 1.	11
Tabla 2. Base de datos de revisiones de literaturas científicas parte 2.	11
Tabla 3. Base de datos de revisiones de literaturas científicas parte 3.	12
Tabla 4. Base de datos de revisiones de literaturas científicas parte 4.	12
Tabla 5. Base de datos de revisiones de literaturas científicas parte 5.	13
Tabla 6. Base de datos de revisiones de literaturas científicas parte 6.	13
Tabla 7. Base de datos de revisiones de literaturas científicas parte 8.	14
Tabla 8. Base de datos de revisiones de literaturas científicas parte 7.	14
Tabla 9. Base de datos de revisiones de literaturas científicas parte 9.	15
Tabla 10. Base de datos de revisiones de literaturas científicas parte 10.	15
Tabla 11. Base de datos de revisiones de literaturas científicas parte 11.	16
Tabla 12. Base de datos de revisiones de literaturas científicas parte 12.	16
Tabla 13. Base de datos de revisiones de literaturas científicas parte 13.	17
Tabla 14. Base de datos de revisiones de literaturas científicas parte 14.	17
Tabla 15. Base de datos de revisiones de literaturas científicas parte 15.	18
Tabla 16. Base de datos de revisiones de literaturas científicas parte 16.	18
Tabla 17. Base de datos de revisiones de literaturas científicas parte 17.	19
Tabla 18. Base de datos de revisiones de literaturas científicas parte 18.	19
Tabla 19. Base de datos de revisiones de literaturas científicas parte 19.	20
Tabla 20. Base de datos de revisiones de literaturas científicas parte 21.	21
Tabla 21. Base de datos de revisiones de literaturas científicas parte 22.	21
Tabla 22. Base de datos de revisiones de literaturas científicas parte 23.	22
Tabla 23. Base de datos de revisiones de literaturas científicas parte 24.	22
Tabla 24. Metodologías de estudios empleados en las investigaciones.	25



ÍNDICE DE FIGURAS

Figura 1. Resultados según buscadores.	24
Figura 2. Procedencia de búsquedas de información.....	24
Figura 3. Procedencia de búsquedas según continentes.	25
Figura 4. Fecha de publicaciones de investigaciones por año.	26
Figura 5. Porcentaje de casos que se evaluarán en la presente investigación.....	26
Figura 6. Tipos de Modelos predictivos usado en las investigaciones.	27



RESUMEN

Hoy en día la deserción es uno de las preocupaciones que recae en cada una de las entidades universitarias, sean públicas o privadas, por lo cual, para poder detectarlo con anticipación se han creado diversos modelos predictivos, esos modelos predictivos darán herramientas a las universidad para la retención de dichos alumnos y que estos culminen la carrera que en una primera instancia decidieron estudiar, las fuentes de información fueron principalmente en Redalyc, Scielo y Google Académico, la elegibilidad de los artículos a investigar se dieron discriminando publicaciones que no sean de habla hispana y de Latinoamérica, así mismo un periodo de máximo 12 años. Mediante este proceso se ha podido identificar modelos predictivos de deserción, así como que tipos de datos y variables son necesarios para poder llevar acabo dichos modelos y de manera adicional que algunas herramientas que de manera conjunta fortalecen dichos modelos. Las limitaciones del presente trabajo fueron principalmente que no se encontraron muchos modelos predictores de deserción

PALABRAS CLAVES: Modelo predictivo, deserción, universitaria, pre grado.



CAPÍTULO I. INTRODUCCIÓN

Para Mori (2012) según su investigación indica que se alcanza una tasa del 17% de deserción estudiantil, por otro lado, Guillermo Sánchez Hernandez & Miguel Barboza-Palomino & Humberto-Castilla-Cabello (2017) su estudio indica que dentro de las Carreras Profesionales de la ciencia de la salud hay un 10% de alumnos que abandonan sus estudios. La deserción es influenciada principalmente por factores económicos, personales- principalmente por la errada vocación profesional que tienen al ingresar- e institucionales.

Su presencia es visible en el escenario educativo latinoamericano, y el Perú no es ajeno a dicho fenómeno; sin embargo, los estudios que indagan a profundidad sus causas y consecuencias resultan escasos: básicamente, se reportan los porcentajes del promedio anual de deserción (Mori, 2012; Heredia et al., 2015).

Es por esto que (Girón y González, 2005; Huesca y Castaño, 2007; Donoso y Schiefelbein, 2007; Díaz, 2008; Abdala et al., 2008; Ariza y Marín, 2009; Medrano et al., 2010; Mateus et al., 2011; Reyes et al., 2012; Piratoba y Barbosa, 2013; Celis et al., 2013) realizaron un análisis predictivo de deserción mediante un análisis estadístico (SPSS) con la información recopilada.

Para Mauricio A. Miranda & Jheser Guzmán (2017) utilizan la técnica minería de datos, esto tipos de clasificadores alcanza niveles estadísticos altos de detección, según sus estos, sea red Bayesiana, árbol de decisión y red neuronal alcanzaría un porcentaje de acierto de 76%, 75% y 83% respectivamente.

Para José Pablo Del Río Arteaga (2018) usar una metodología CRISP-DM y juntándolo con algoritmos de clasificación como bosques aleatorios alcanzó un nivel del 96% de alumnos no desertores.

Por consiguiente, la pregunta a realizarnos será ¿Cuáles son los modelos predictivos que han estimado la deserción de estudiantes durante el periodo de años 2008 y 2018? El objetivo del presente trabajo será el describir los modelos predictivos universitarios que ubican a alumnos con un potencial alto de deserción. Analizaremos los estudios teóricos y empíricos entre los años 2008 – 2018 de Latinoamérica. La metodología utilizada fue de revisión sistemática de la literatura científica.



CAPÍTULO II. METODOLOGÍA

El presente trabajo se ha realizado mediante la revisión sistemática de la literatura científica. Para Mikhail Benet Rodríguez, Sandra Liliana Zafra, Sandra Patricia Quintero Ortega (2015) no hace reflexionar sobre la importancia de la revisión sistemática de la literatura científica, dado que es de necesidad de todo investigador revisar literaturas nacionales e internacionales con temas de interés. Estas investigaciones a través del análisis documental contribuyen al desarrollo de la ciencia e instituciones.

Según Jorge Iván Pérez Rave (2012) en su libro *Revisión Sistemática de Literatura en Ingeniería*, nos indica que la literatura científica se encuentra compuesta por 4 procesos (identificar, describir, profundizar, divulgar), estas nos permitirán deducir las condiciones actuales y cuáles serán los desafíos a posteriori, esto permitirá darnos reproducibilidad de los resultados, el cual garantiza la eficacia en la identificación de las investigaciones más posicionadas y para que de este modo ayude a los investigadores o estudiantes en formación.

Como se comentó al inicio de la presente investigación, el presente objetivo es encontrar cuáles son los modelos predictivos que se utilizan en diversas entidades educativas universitarias para poder encontrar a los alumnos con mayor potencial de deserción. La estrategia para poder corroborar que no se hayan duplicado las investigaciones, se ha tenido que leer y visualizar cada una de las investigaciones para evitar este inconveniente.



BASE DE DATOS DE REVISIONES DE LITERATURAS CIENTÍFICAS

Sumatoria	Keyword	BD	Accesibilidad	APA	Autores	Título
1	Modelo predictivo, deserción, universitaria, pre grado	Redalyc	Abierto	Apaza, E., & Huamán, F. (2012). Factores determinantes que inciden en la deserción de los estudiantes universitarios. <i>Apuntes Universitarios. Revista de Investigación</i> , (1), 77-86.	Apaza, Elmer; Huamán, Francisco	Factores determinantes que inciden en la deserción de los estudiantes universitarios
2	Modelo predictivo, deserción, universitaria, pre grado	Redalyc	Abierto	Hernero, V., & Medina, A., & Aylón, S., & Escobar, G. (2013). Aplicación de un modelo de duración en programas de prevención de deserción universitaria. <i>REDIE. Revista Electrónica de Investigación Educativa</i> , 15 (3), 38-52.	Hernero, Verónica; Medina, Aldo; Aylón, Silvia; Escobar, Gabriel	Aplicación de un modelo de duración en programas de prevención de deserción universitaria
3	Modelo predictivo, deserción, universitaria, pre grado	Redalyc	Abierto	SALCEDO, R., & CORREA, R. (2008). SICOMP: SIMULADOR PARA CONTROL PREDICTIVO BASADO EN MODELOS. <i>Dyna</i> , 75 (169), 89-98.	ROBIN SALCEDO; RODRIGO CORREA	SIMULADOR PARA CONTROL PREDICTIVO BASADO EN MODELOS
4	Modelo predictivo, deserción, universitaria, pre grado	Redalyc	Abierto	BARAHONA URBINA, P., & VERES FERRER, E., & ALUAGA PRIETO, V. (2016). DESERCIÓN ACADÉMICA DE LA UNIVERSIDAD DE ATACAMA, CHILE. COMUNICACIÓN. <i>Revista de Investigación en Comunicación y Desarrollo</i> , 7 (2), 27-37.	PLANCK BARAHONA URBINA; ERNESTO VERES FERRER; VERÓNICA ALUAGA PRIETO	DESERCIÓN ACADÉMICA DE LA UNIVERSIDAD DE ATACAMA, CHILE

Tabla 1. Base de datos de revisiones de literaturas científicas parte 1.

Abstract	Metodologías
Este artículo aborda el tema de la deserción de estudiantes en la educación superior universitaria desde una perspectiva conceptual. Se analizan las distinciones que son posibles de efectuar con respecto a este término, con el propósito de poder adoptar y entender ampliamente las definiciones para futuros trabajos sobre factores que influyen en la deserción universitaria. Además de haber recolectado una gran variedad de conceptos psicológicos, económicos, sociológicos, organizacionales y de las interacciones entre el estudiante y la institución. Dichos factores son predictivos para el abandono estudiantil, dado que habría que dar más cuidado a algunas variables que son más críticas para la toma de decisiones en los estudiantes. Las prácticas institucionales vinculadas con la prevención de la deserción de una carrera universitaria requieren cada vez más de instrumentos validados que permitan anticipar tal comportamiento. En este sentido han demostrado utilidad decisiva distintos modelos estadísticos generados a partir de información referida a los propios estudiantes, sus hogares y su desempeño académico, entre otros determinantes. Este estudio pretende mostrar la importancia de una serie de determinantes exploradas en otros estudios: el objetivo principal es aplicar un modelo predictivo del riesgo de deserción de estudiantes universitarios a fin de generar resultados de manera temprana y progresivamente más eficaces. El trabajo exhibe la utilidad de los modelos de duración en una muestra de estudiantes presenciales, y la capacidad anticipatoria de los comportamientos de permanencia/deserción en el tiempo, a través de estimaciones de riesgo con un modelo de Cox en cuatro momentos de los primeros meses de universidad. En este artículo se describe la herramienta computacional SICOMP desarrollada con el fin de poder disponer de un programa para simulación de control predictivo basado en modelos. Si bien existen en el mercado algunos programas dedicados al control predictivo, ellos son en su mayoría costosos y de estructura cerrada. SICOMP de otro lado tiene una estructura modular e incluye además ventanas donde se puede simular las estrategias clásicas de control, con el fin de compararlos bajo criterios de estabilidad y robustez en diversos modelos de plantas de procesos. Este trabajo tiene doble propósito, el primero de ellos es determinar los factores asociados al rendimiento académico de los alumnos del primer año lectivo 2014 de las carreras de Ingeniería de la Universidad de Atacama; el segundo, determinar aquellas variables determinantes de la deserción académica. De acuerdo a la naturaleza de los datos se hizo uso del modelo de regresión lineal múltiple, para determinar los factores asociados al rendimiento académico. En la segunda parte se utilizó un modelo de Regresión Logística para determinar aquellos factores que han incidido en la tasa de deserción. Para medir el rendimiento académico se empleó el ratio créditos aprobados sobre créditos inscritos y como variables exógenas, distintos factores de carácter socioeconómico, institucional, académico e individual. Los resultados pusieron de manifiesto que las variables explicativas del rendimiento académico están relacionadas con las notas de ingreso a la Universidad, la asistencia a clase y el tipo de tipología del establecimiento de procedencia. En cuanto al segundo de los objetivos, se encontró que las variables asociadas a la deserción académica tienen que ver con las características de inscripción a la universidad, esto es, créditos inscritos como medida del esfuerzo inicial del estudiante y el rendimiento académico.	Empírico cualitativo tipo de estudio de caso, evaluación de carácter documental de la plataforma, no tiene variables, pero se indica que en este estudio se evaluó la calidad del MOOC metodología de riesgo proporcional de Prince y Gloedier

Tabla 2. Base de datos de revisiones de literaturas científicas parte 2.

¿En qué revistas se publicaron estos estudios?	País	Año	¿Qué tipo de estudios se realizaron?
Apuntes Universitarios	Perú	2012	Cualitativos y cuantitativos.
REDIE. Revista Electrónica de Investigación Educativa	México	2013	Cualitativos y cuantitativos.
Dyna	Colombia	2008	Cuantitativos y cualitativos
COMUNI@CCION: Revista de Investigación en Comunicación y Desarrollo	Chile	2016	Mixto

Tabla 3. Base de datos de revisiones de literaturas científicas parte 3.

¿Qué variables se analizaron?	¿Cómo se realizó la medición? (instrumento)	exclusión	exclusión normalizada
Variables explicativas	Datos de los alumnos de la universidad, varios.	Principalmente habla sobre las variables de deserción, más no de algún modelo predictivo.	Excluido
Variables explicativas	Modelo de Cox	Se excluye dado que es de norteamérica	Excluido
Multivariables	SICOMPC	Se excluye dado que se trata de control a nivel industrial.	Excluido
Variables predictoras	Encuestas, datos del departamento de computación de la Universidad de Atacama.	No se excluye	No se excluye

Tabla 4. Base de datos de revisiones de literaturas científicas parte 4.

Sumatoria	Keyword	BD	Asociabilidad	APA	Revisión	Título
5	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Sanabria, H. (2002). Deserción en estudiantes de enfermería en cuatro universidades del Perú. <i>Anales de la Facultad de Medicina</i> , 63(4), 301-311.	HEPNAU/ SANABRIA	Deserción en estudiantes de enfermería en cuatro universidades del Perú
6	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Saldana Villa, M., & Baranga, O. (2010). Adaptación del modelo de deserción universitaria de Tinto a la Universidad Católica de la Santísima Concepción, Chile. <i>Revista de Ciencias Sociales (Val.)</i> , 16(14), 616-628.	Saldana Villa, Magdalena; Baranga, Omar A.	Adaptación del modelo de deserción universitaria de Tinto a la Universidad Católica de la Santísima Concepción, Chile
7	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Belanga Silvente, V., & Figueroa Gato, M. & Ponce Fariña, E. (2018). MODELO PREDICTIVO DE PERSISTENCIA UNIVERSITARIA: ALUMNADO CON BECA SALARIO. <i>Educación XXI</i> , 21(1), 203-230.	Belanga Silvente, Vanessa; Figueroa Gato, Myriam Pilar; Ponce Fariña, Enwat	MODELO PREDICTIVO DE PERSISTENCIA UNIVERSITARIA: ALUMNADO CON BECA SALARIO

Tabla 5. Base de datos de revisiones de literaturas científicas parte 5.

Autores	Metodología
Determinar los factores de riesgo asociados a la interrupción de los estudios de enfermería en algunas universidades del Perú, así como determinar la magnitud de la deserción según factor de riesgo. DISEÑO: Tipo caso-control, durante 1999-2000. MATERIAL Y MÉTODOS: Se estudió los siguientes factores de riesgo: salud, económicos, personales, familiares, vocacionales, laborales, judiciales-politiales, académico y de adaptación a la vida universitaria. La muestra estuvo conformada por 98 casos distribuidos como sigue: 24 estudiantes de las ciudades de Huacho e Iquitos y 20 de las ciudades de Lima y Trujillo. La muestra no incluyó a estudiantes de la UNHSM, por no contar con la autorización correspondiente. Los controles estuvieron conformados por 65 alumnos no desercionados de las mismas universidades. RESULTADOS: El factor de riesgo de deserción que tuvo la mayor fuerza de asociación fue el factor vocacional, seguido del factor económico y del factor académico, acorde al odds ratio y sus límites inferiores de intervalo de confianza 95% superiores a 1. Un 42% de estudiantes presentó a la falta vocación como factor de riesgo de deserción, seguido del factor económico con 36%. CONCLUSIONES: Existe una alta asociación entre los factores vocacional y económico, así como, una leve a moderada asociación del factor académico-rendimiento con la deserción de los estudiantes de enfermería de las universidades estudiadas. En esta investigación se adaptó el modelo de Vincent Tinto para analizar la deserción de los alumnos de Ingeniería Civil en la Universidad Católica de la Santísima Concepción, Chile. El objetivo de este estudio es identificar qué factores afectan la permanencia de los estudiantes en la carrera escogida, considerando características propias de los alumnos, el compromiso con la institución y el rendimiento académico al ingresar a la universidad. La metodología utilizada en el análisis de eventos históricos y se modela a través de una regresión logística binaria, construyendo un modelo que predice correctamente el 50,31% de los casos. Las variables que mejor explican la deserción son los ingresos familiares y el rendimiento académico de los alumnos; sin embargo, se plantea la necesidad de repensar la relación entre ingresos y deserción a la luz de los resultados de este estudio. Se concluye que el modelo de Tinto y su adaptación a la realidad chilena son un esquema conceptual válido para examinar el fenómeno de la deserción. La dimensión social de la Educación Superior constituye uno de los elementos fundamentales de las políticas universitarias a nivel internacional en estos inicios del siglo XXI. Entre las medidas de retención, las políticas de becas y ayudas al estudio favorecen la equidad ya que garantiza una renta a las familias con menor nivel de recursos. En España, dentro del marco de la Estrategia 2015, la implantación de las becas salario pretendió garantizar el acceso y la persistencia a los estudios universitarios a aquellos grupos sociales más desfavorecidos económicamente. La presente investigación utiliza un modelo de regresión logística dirigido a buscar los factores determinantes de la permanencia del alumnado becado. El estudio se ha realizado con un total de 642 estudiantes de grado con beca salario de la cohorte 2010-11 de la Universidad de Barcelona. Con base a los resultados se puede afirmar que la probabilidad de que un o una estudiante becado de la población objeto de estudio presente persistencia, aumenta con el incremento de la tasa de persistencia y con un alto porcentaje de asignaturas aprobadas. Los resultados estimados parecen indicar que la beca proporciona mayor dedicación al estudio y permite compensar las dificultades del alumnado procedente de familias con bajos ingresos. Y por lo tanto se constituye en un instrumento potente para garantizar la equidad social.	Análisis de eventos históricos descriptivo-comparativo

Tabla 6. Base de datos de revisiones de literaturas científicas parte 6.

¿En qué revistas se publicaron estos estudios?	País	Año	¿Qué tipo de estudios se realizaron?
Anales de la Facultad de Medicina	Perú	2002	Caso - control
Revista de Ciencias Sociales (Ve)	Venezuela	2010	Cualitativos y cuantitativos.
Educación XX1	España	2018	Cualitativos y cuantitativos.

Tabla 8. Base de datos de revisiones de literaturas científicas parte 7.

¿Qué variables se analizaron?	¿Cómo se realizó la medición? (instrumento)	exclusión	exclusión normalizada
Variables explicativas	Datos de los alumnos de la universidad, varios.	Se excluye, por el tiempo	Excluido
Variables predictoras	Datos de los alumnos de la universidad, varios.	No se excluye	No se excluye
Variables explicativas y descriptivas.	Datos de los alumnos de la universidad, varios.	Excluido dado que son carreras a distancia y de procedencia Europea.	Excluido

Tabla 7. Base de datos de revisiones de literaturas científicas parte 8.

Sumatoria	Keyword	ID	Accesibilidad	APA	Autores	Título
8	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Londoño Ardila, L. (2013). Factores de riesgo presentes en la deserción estudiantil en la Corporación Universitaria Lasallista. <i>Revista Virtual Universidad Católica del Norte</i> , 13(1), 160-194.	Luis Felipe Londoño Ardila	Factores de riesgo presentes en la deserción estudiantil en la Corporación Universitaria Lasallista
9	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Eckert, K., & Suñéaga, R. (2015). Análisis de Deserción-Permanencia de Estudiantes Universitarios Utilizando Técnica de Clasificación en Minería de Datos. <i>Formación Universitaria</i> , 8(15), 3-12.	Eckert, Katina B.; Suñéaga, Roberto	Análisis de Deserción-Permanencia de Estudiantes Universitarios Utilizando Técnica de Clasificación en Minería de Datos
10	Modelo predictivo, deserción, universitaria, pre grado.	Google Académico	Abierto	Cogua Bohórquez M. E. (2016). Modelo predictivo de deserción en un sistema de educación a distancia. Trabajo de Grado. Universidad Católica de Colombia. Facultad de Psicología. Maestría en Psicología. Bogotá, Colombia.	Cogua Bohórquez, William Eneider	MODELO PREDICTIVO DE DESERCIÓN UNIVERSITARIA EN UN SISTEMA DE EDUCACIÓN A DISTANCIA

Tabla 9. Base de datos de revisiones de literaturas científicas parte 9.

Autores	Resumen de la literatura
	El propósito de este estudio fue describir los factores de riesgo personales, académicos, institucionales y socioeconómicos asociados a la deserción estudiantil en la Corporación Universitaria Lasallista (CUL). La muestra estuvo conformada por 281 estudiantes activos en diferentes programas de la institución en el año 2010 y 31 estudiantes que habían desertado en el 2003. Se construyeron dos escalas que indagaban por factores que podían incidir en la deserción en los estudiantes activos y los que habían interrumpido sus estudios en la universidad. El análisis descriptivo utilizado consistió en la distribución de frecuencias y la obtención de medidas de tendencia central para encontrar aquellos factores que según los participantes se relacionaban con la deserción. Los resultados mostraron que los factores de riesgo para la deserción académica son predominantemente socioeconómicos, posteriormente se encuentran los institucionales y con una baja distribución porcentual están los factores académicos y Se analiza información académica con el objetivo de identificar factores que influyen sobre la deserción de los estudiantes de la carrera de Ingeniería en Informática de la Universidad Gaspar Dachary en Argentina, mediante la aplicación de una técnica de minería de datos. La fuente de datos contiene información proporcionada al ingreso (personales y antecedentes educativos) y la que se genera durante el periodo de estudios. Se realiza la selección y depuración de datos, utilizando diferentes criterios de representación y aplicación de algoritmos de clasificación como árboles de decisión, redes bayesianas y reglas. Se identifica como variables influyentes en la deserción, asignaturas aprobadas, cantidad y resultado de asignaturas cursadas, procedencia y edad de ingreso del estudiante. Mediante este proceso fue posible identificar los atributos que caracterizan a los casos de deserción y su relación con el desempeño académico, especialmente en el primer año de la carrera.
	Investigación que buscó determinar qué variables explican la deserción estudiantil en un sistema de educación a distancia, a partir de la construcción de un modelo de predicción que incluye siete categorías de variables.

Tabla 10. Base de datos de revisiones de literaturas científicas parte 10.

¿En qué revistas se publicaron estos estudios?	País	Año	¿Qué tipo de estudios se realizaron?
Revista Virtual Universidad Católica del Norte	Colombia	2013	Cualitativos y cuantitativos.
Formación Universitaria	Chile	2015	Cualitativos y cuantitativos.
ADB. Maestría en Psicología	Colombia	2016	-

Tabla 11. Base de datos de revisiones de literaturas científicas parte 11.

¿Qué variables se analizaron?	¿Cómo se realizó la medición? (instrumento)	exclusión	exclusión normalizada
Variables explicativas	Cuestionarios hetero-aplicados.	Excluido dado que no presenta un modelo predictivo.	Excluido
Variables explicativas y descriptivas.	Técnica de Clasificación en Minería de datos	No se excluye	No se excluye
-	-	Se encuentra cerrado la publicación.	Excluido

Tabla 12. Base de datos de revisiones de literaturas científicas parte 12.

Sumatoria	Keyword	BD	Accesibilidad	APA	Autores	Título
11	Modelo predictivo, deserción, universitaria, pre grado.	Google Académico	Abierto	Gálvez Chambiella, M., & Flores Cornejo, K. (2018). Modelo predictivo de deserción universitaria de la carrera de Ingeniería Informática en la Universidad Ricardo Palma. Retrieved from: http://ojs.biblioteca.up.edu.pe/handle/123456789/12345	Gálvez Chambiella, Melfora Beatriz; Flores Cornejo, Katherine Brigitte	Modelo predictivo de deserción universitaria de la carrera de Ingeniería Informática en la Universidad Ricardo Palma
12	Modelo predictivo, deserción, universitaria, pre grado.	Google Académico	Abierto	Yegry Kaina Anaya Toralosa, Edwin Benavides Avendaño, Diana Judith Heredia Vozainchob. (2018). Modelo predictivo de deserción estudiantil utilizando técnicas de minería de datos. 22(03)2018, de Universidad Simón Bolívar, Barranquilla, Colombia.	Yegry Kaina Anaya Toralosa, Edwin Benavides Avendaño, Diana Judith Heredia Vozainchob	Modelo predictivo de deserción estudiantil utilizando técnicas de minería de datos

Tabla 13. Base de datos de revisiones de literaturas científicas parte 13.

Abs
La deserción universitaria se ha convertido en un problema importante a ser investigado. La tasa de deserción ha llegado a constituir uno de los principales indicadores de eficiencia interna dentro de cualquier institución de educación. Investigar las causas de la deserción con metodologías adecuadas que permitan predecir esta, contribuye a la toma de decisiones dentro de la gestión de la unidad académica. El objetivo del presente proyecto consiste en investigar y proponer una metodología que permita identificar en forma automática a los estudiantes con mayor riesgo de deserción de las carreras de Ingeniería Informática en la Universidad Ricardo Palma. Para la implementación de este proyecto se adoptó la metodología CRISP-DM que estructura el proceso de minería de datos en seis fases, que interactúan entre ellas de forma iterativa. Se evaluaron los modelos de Árboles de decisión y Clustering para analizar el comportamiento de los estudiantes, evaluando factores como el rendimiento del alumno, condición social y aspectos socioeconómicos. La exactitud de los modelos es calculada a partir de la información que brindó la Oficina Central de Informática y Cómputo de la Universidad Ricardo Palma, en la cual se realizó una transformación y simulación de algunas variables para mayor efectividad del modelo. The college desertion has become an important problem to be investigated. The desertion rate has come to be one of the main indicators of internal efficiency in any educational institution. The research of the causes of desertion with appropriate methodologies to predict, contributes to decision making in the academic management. The objective of this project is to investigate and propose a methodology to automatically identify students at high risk of dropping out of the Computer Engineering career at the Ricardo Palma University. To implement this project, we used the CRISP-DM methodology to structure the data mining process into six phases, which interact with each other iteratively. It was applied the models of decision trees and clustering to analyze the behavior of the students, evaluating factors such as student performance, social status and socio-economic aspects. The accuracy of the models is calculated from the information provided the Central Office of Information and Computation at the Ricardo Palma University, in which it was applied a transformation and simulation of some variables for greater effectiveness and performance of the model. La minería de datos o Data Mining permite descubrir información oculta en grandes cantidades de datos, información que por procedimientos tradicionales es muy difícil de visualizar. Esta rama de la computación permite manejar y clasificar grandes cantidades de datos, para lo cual se utilizan una gran variedad de técnicas, entre las que se encuentran los árboles de decisión C4.5 y el ID3 que han demostrado ser muy eficientes para casos específicos de predicción; este tipo de técnicas generan árboles que, de acuerdo a la complejidad del tema de estudio, pueden ser muy variables: se pueden obtener árboles con muchos nodos y hojas en el caso de ID3 y árboles más pequeños si utilizamos C4.5. Este artículo muestra la construcción de un modelo predictivo de deserción estudiantil, caracterizando a los estudiantes de la Universidad Simón Bolívar con el objetivo de poder predecir la probabilidad de deserción de los estudiantes; dicho modelo demostró el desempeño de los algoritmos presentados para clasificar datos bajo contextos variables y la precisión de uno con respecto al otro. Para la creación del modelo se utilizó la herramienta WEKA que permite de forma muy eficiente el procesamiento y clasificación de los datos con resultados satisfactorios.

Tabla 14. Base de datos de revisiones de literaturas científicas parte 14.

¿En qué revistas se publicaron estos estudios?	País	Año	¿Qué tipo de estudios se realizaron?
Repositorio Digital - URP	Perú	2015	Cualitativos y cuantitativos.
Documentos Red clara	Colombia	2014	Cualitativos y cuantitativos.

Tabla 15. Base de datos de revisiones de literaturas científicas parte 15.

¿Qué variables se analizaron?	¿Cómo se realizó la medición? (instrumento)	exclusión	exclusión normalizada
Variables explicativas	Vincent Tinto y árboles de deserción. CRISP-DM	No se excluye	No se excluye
Variables explicativas y descriptivas.	Data mining	No se excluye	No se excluye

Tabla 16. Base de datos de revisiones de literaturas científicas parte 16.

Sumatoria	Keyword	BD	Accesibilidad	APA	Autores	Tema
13	Modelo predictivo, deserción, universitaria, pre grado.	Google Académico	Abierto	Pablo Barcoel Baeza & Sebastian Rios Perez & Ma. Angélica Pinnighoff Jurewicz. (2020) MODELO PARA LA AUTOMATIZACIÓN DEL PROCESO DE DETERMINACIÓN DE RIESGO DE DESERCIÓN EN ALUMNOS UNIVERSITARIOS. Tesis. Santiago de Chile, Chile.	Pablo Barcoel Baeza & Sebastian Rios Perez & Ma. Angélica Pinnighoff Jurewicz	MODELO PARA LA AUTOMATIZACIÓN DEL PROCESO DE DETERMINACIÓN DE RIESGO DE DESERCIÓN EN ALUMNOS UNIVERSITARIOS
14	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Bernardo, A. et al. (2015). Predicción del abandono universitario: variables explicativas y medidas de prevención. Revista Fuentes, 16, Junio, pp. 63-84.	Ana Balón, Bernardo Guzmán, Rebecca Cerezo, Mercedes Lito J, Rodríguez Muñoz, José Carlos Nájera, Pérez Elías, Tzuc Henao, María Esteban García	Factores determinantes que inciden en la deserción de los estudiantes universitarios
15	Modelo predictivo, deserción, universitaria, pre grado.	Google Académico	Abierto	Acounana, K. (2013). Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos. Pensamiento Americano, 41-51.	Kamagase Acounana	Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos.

Tabla 17. Base de datos de revisiones de literaturas científicas parte 17.

Tema	Resumen de la revisión
	Para la implementación de este proyecto se adoptó la metodología CRISP-DM que estructura el proceso de minería de datos en seis fases, que interactúan entre ellas de forma iterativa. Se aplicaron los modelos de Redes Neuronales, Árboles de decisión y Clúster K-medias para analizar el comportamiento de los estudiantes, evaluando factores como el puntaje promedio obtenido en la Prueba de Selección Universitaria (PSU), el promedio de notas obtenido en la enseñanza media, La edad a la fecha de ingreso a la institución y el género de los estudiantes.
	El abandono de los estudios universitarios es un problema cuyos costos son altos tanto para el individuo como para la sociedad. Es por ello que la prevención del mismo es fundamental y cobra especial relevancia en el actual contexto de crisis económica. Diversos autores han desarrollado investigaciones con el objetivo de establecer modelos predictivos de este fenómeno (Castaño, Gallón, Gómez y Viquez, 2004; Trevián, Beltrán y Corallo, 2009; Guldenharth, Coxa y Saino, 2015; Sánchez, 2014). En este artículo se analizan dichos trabajos, identificando las ventajas y desventajas de las metodologías más utilizadas, análisis correlacionales, análisis de regresión logística, análisis de supervivencia y minería de datos. La investigación cuyos resultados aquí se exponen, aplica la primera de las metodologías mencionadas, a fin de comprobar –en lo que respecta al fenómeno del abandono– el valor predictivo de las variables rendimiento académico previo, fecha de matriculación, rendimiento en primer curso de universidad y asistencia a clase. Los resultados confirman la relación de dichas variables con el fenómeno estudiado. Dichos resultados son consistentes con los obtenidos por diversos autores a lo largo del tiempo, y en base a ellos se proponen dos tipos de medidas: por un lado, acciones encaminadas a facilitar el diagnóstico respecto al problema del abandono, y por otro lado, medidas encaminadas a su prevención.
	El presente artículo muestra los resultados del análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad de Ingeniería de Sistemas, con técnicas de minería de Datos. Se utilizó la herramienta Weka agrupando las causas de la deserción en 5 variables que son: Pérdida de semestre, Dificultad financiera, Ingreso al mercado laboral Otro intereses avisan al estudiante, Indeterminado. La muestra era de 707 sujetos entre los periodos 2007-2012. Se concluye que la causa mayor de la deserción es el factor indeterminado.

Tabla 18. Base de datos de revisiones de literaturas científicas parte 18.

¿En qué revistas se publicaron estos estudios?	País	Año	¿Qué tipo de estudios se realizaron?
Repositorio Universidad de Chile	Chile	2012	Cualitativos y cuantitativos.
Revista Fuentes	España	2015	Cualitativos y cuantitativos.
Revista Pensamiento Americano	Colombia	2013	Cualitativos y cuantitativos.

Tabla 19. Base de datos de revisiones de literaturas científicas parte 19.

¿Qué variables se analizaron?	¿Cómo se realizó la medición? (instrumento)	exclusión	exclusión normalizada
Variables Descriptivas	Data Mining y programación	Se excluye dado que es de norteamérica	Excluido
Variables explicativas	Datos de los alumnos de la universidad, varios.	Principalmente habla sobre las variables de deserción, más no de algún modelo predictivo.	Excluido
Variables explicativas	Minería de datos y herramienta WEKA	No se excluye	No se excluye

Tabla 20. Base de datos de revisiones de literaturas científicas parte 20.

Sumario	Keyword	BD	Disponibilidad	APA	Autores	Título
16	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Herniquez-Cabezas, N. & Escobar-Rifo, D. (2018). CONSTRUCCIÓN DE UN MODELO DE ALERTA TEMPRANA PARA LA DETECCIÓN DE ESTUDIANTES EN RIESGO DE DESERCIÓN DE LA UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACIÓN. Revista Mexicana de Investigación	Herniquez-Cabezas, Nanaia; Escobar-Rifo, Diego	CONSTRUCCIÓN DE UN MODELO DE ALERTA TEMPRANA PARA LA DETECCIÓN DE ESTUDIANTES EN RIESGO DE DESERCIÓN DE LA UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACIÓN
17	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	Miranda, M., & Guzmán, J. (2017). Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos. Formación Universitaria, 10 (3), 61-68.	Mauricio A. Miranda y Jhesar Guzmán	Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos
18	Modelo predictivo, deserción, universitaria, pre grado.	Redalyc	Abierto	José Pablo del Río Arteaga. (2018). MODELO PREDICTIVO PARA LA RETENCIÓN DE ESTUDIANTES EN PRIMEROS AÑOS. Tesis de Universidad de la Rioja	DEL RÍO ARTEAGA, JOSÉ PABLO	MODELO PREDICTIVO PARA LA RETENCIÓN DE ESTUDIANTES EN PRIMEROS AÑOS
19	Modelo predictivo, deserción, universitaria, pre grado.	Scielo	Abierto	Panico E. Ramirez, Elizabeth E. Grandón. (2017). Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados. Tesis de Universidad Católica del Norte.	Panico E. Ramirez, Elizabeth E. Grandón	Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados
20	Modelo predictivo, deserción, universitaria, pre grado.	Google Académico	Abierto	N. E. Rodríguez Maya, A. J. Jiménez Alfaro, L. A. Reyes Hernández, B. A. Suarez Caranza and J. K. Ruiz Garduno, "Data mining: a scholar dropout predictive model," 2017 IEEE Mexican Humanitarian Technology Conference (MHTC), Puebla, 2017, pp. 89-93.	Noel Enrique Rodríguez Maya, Abraham Jorge Jimenez Alfaro, Luis Angel Reyes Hernandez, Brian Alison	Data mining: a scholar dropout predictive model

Tabla 21. Base de datos de revisiones de literaturas científicas parte 21.

Autores	Resumen/objetivo
El presente artículo analiza qué variables se relacionan con el manejo de habilidades requeridas al momento de ingresar a la Universidad Metropolitana de Ciencias de la Educación (UMCE); el propósito es establecer índices para la detección de estudiantes que requieren apoyo durante el primer año académico, que es de alta deserción. El modelamiento se realizó mediante regresión logística con análisis de curvas ROC. Dentro de los resultados obtenidos, las variables que permiten pronosticar deficiencias en habilidades de lenguaje y matemática fueron: puntajes de la prueba de selección universitaria en ambas materias, promedio de notas de la enseñanza media, género y edad de ingreso a la UMCE. Con estas variables se estableció una clasificación de los estudiantes (puntos de corte y probabilidades) para establecer distintas estrategias de apoyo académico.	Muestreos aleatorios estratificado.
Se realiza un estudio para determinar cuáles son y cuáles es la importancia de las variables que llevan a un estudiante a abandonar sus estudios universitarios, usando técnicas de minería de datos. La deserción de los estudiantes de educación superior genera una serie de inconvenientes que afectan a los estudiantes y las universidades. Los resultados obtenidos a partir de los datos proporcionados por las carreras de Ingeniería de la Universidad Católica del Norte en Antofagasta y Coquimbo (Chile) determinan que las variables que mejor explican la deserción de un estudiante son: las razones socioeconómicas y el puntaje de ingreso a la universidad (PSU). Según el árbol de decisión construido se concluye que la retención se sitúa en un 78,3%. La calidad de los clasificadores permite asegurar que sus predicciones son correctas, con niveles estadísticos de curva ROC de 76%, 75% y 63% de acierto para los clasificadores de red bayesiana, árbol de decisión y red neuronal respectivamente.	Data Mining
Se crean modelos capaces de predecir la deserción de estudiantes de Ingeniería Civil de la Universidad Técnica Federico Santa María en el primer y segundo año académico. Se utilizó la información de estudiantes referente al pre ingreso a la universidad, y la información académica del estudiante hasta el término del primer semestre. Los mejores modelos construidos se lograron con los algoritmos random forest, para la predicción de la deserción en el primer año académico y gradient boosting para la predicción de la deserción en el segundo año académico, con valores del 82,4% y 82,3% para la métrica F1 score respectivamente.	Data Mining
El objetivo de este estudio es presentar una clasificación basada en árboles de decisión (CBAD) con parámetros optimizados para predecir la deserción de los estudiantes universitarios. El estudio analiza 5269 casos de estudiantes pertenecientes a una universidad pública chilena. Para la técnica CBAD se optimizaron los parámetros para mejorar la predicción usando el software RapidMiner. El resultado de la aplicación de esta técnica con parámetros optimizados logró una razón de precisión de un 87,27%. Se concluye que el uso de técnicas de CBAD con optimización de parámetros resulta en una mejor precisión en comparación a otras investigaciones con un número similar de datos.	Árboles de deserción: CBAD
The Scholar Dropout (SD) phenomenon in universities has been increasing in the past years, having repercussions in social, economic, and academic aspects among others. There are different factors that affect students to leave their studies and vary according to the place where the action takes place. Data Mining (DM) is a tool that helps to identify hidden patterns through the search of patterns into the data, for example, the creation of models to describe historical data. This research presents a SD predictive model in universities, which is based in a methodology based on DM using as study case, information of the generational cohorts 2010-2015 and 2011-2016 from the Instituto Tecnológico de Zihuatcero. The results showed a predictive model with a precision above 85%. It can be established in other universities with the necessary and pertinent adjustments.	Data Mining

Tabla 22. Base de datos de revisiones de literaturas científicas parte 22.

¿En qué institución se publicaron estos estudios?	País	Año	¿Qué tipo de estudios se realizaron?
Revista Mexicana de Investigación Educativa	México	2016	Cuantitativa
Centro de Información Tecnológica Chile	Chile	2017	Datos Cuantitativos
Repositorio Universidad Técnica Federico Santa María	Chile	2016	Cualitativos y cuantitativos
Scientific electronic library online	Chile	2015	Cualitativos y cuantitativos
2017 IEEE Mexican Humanitarian Technology Conference (MHTC)	México	2017	S/D

Tabla 23. Base de datos de revisiones de literaturas científicas parte 23.

¿Qué variables se analizaron?	¿Cómo se realizó la medición? (instrumento)	exclusión	exclusión normalizada
Variables descriptiva inferencial	Regresión logística multivariado	Toma carreras como Música y educación, no tan centrado en las carreras a conocer.	No se excluye
Variables descriptivas	Minería de datos, árboles de decisión, Bayes, redes naturales.	No se excluye	No se excluye
Variables explicativas y descriptivas.	Minería de datos, algoritmos Random Forest, tgradient y boosting	No se excluye	No se excluye
Variables explicativas y descriptivas.	Arboles de deserción. CBAD.	No se excluye	No se excluye
S/D	Data mining	Se excluye, dado que está en inglés.	Excluido

Tabla 24. Base de datos de revisiones de literaturas científicas parte 24.

CAPÍTULO III. RESULTADOS

A continuación, se presentará las siguientes imágenes con más detalle sobre cuál ha sido la forma de poder escoger los temas más relevantes para el presente estudio. Teniendo en cuenta lo siguiente; el tiempo de antigüedad de los documentos *Figura 4*, los países de procedencia *Figura 2*, así como el continente *Figura 3* y la metodología realizada en los estudios *Tabla 24*. Teniendo toda esta información, llegamos a los motivos de inclusión y exclusión de las investigaciones que serán las siguientes: serán 10 estudios que se tomarán en cuenta; serán de un periodo máximo de 12 años de antigüedad, de variables cualitativas y cuantitativas, sólo a nivel Sudamérica y según los motores de búsqueda o de revistas como Scielo, Redalyc y Google académico *Figura 1*. Finalizando, en la *Figura 6*. luego de haber encontrado las investigaciones que se revisarán, ya habiendo hecho los filtros mencionados líneas arriba, nos damos con la información final de los tipos de modelos usados por las investigaciones recabadas en el presente estudio.

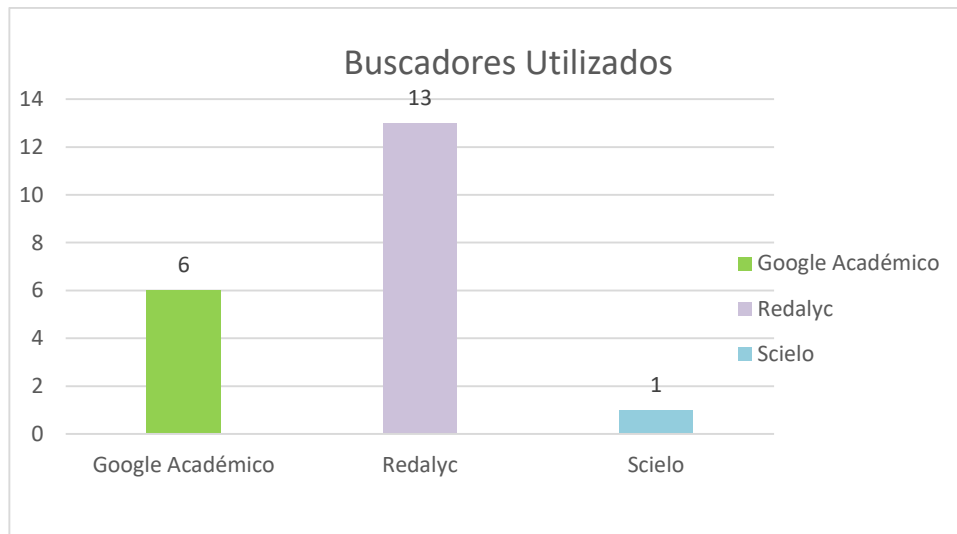


Figura 1. Resultados según buscadores.

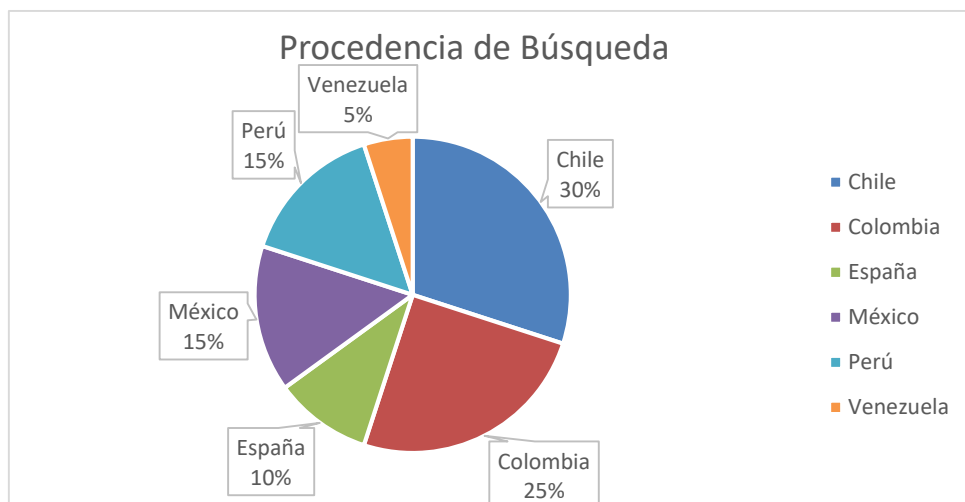


Figura 2. Procedencia de búsquedas de información.

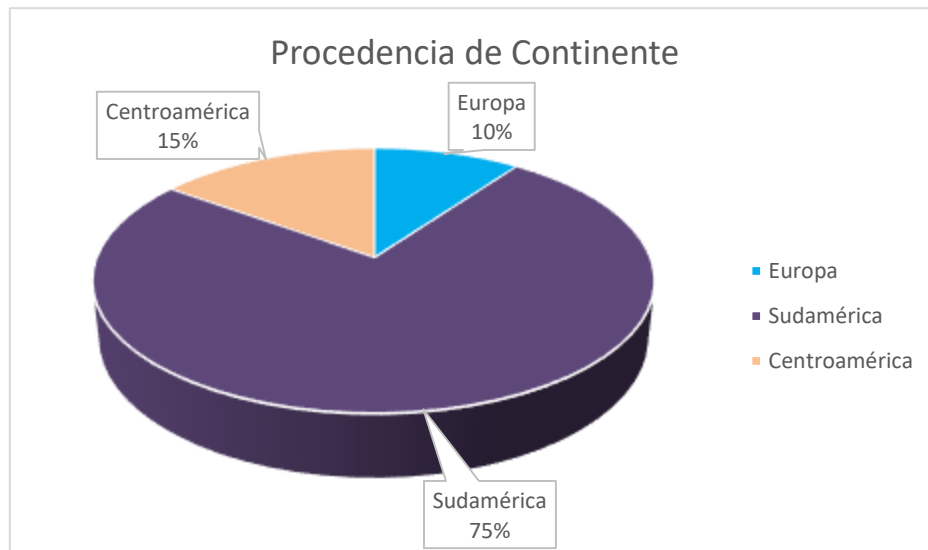


Figura 3. Procedencia de búsquedas según continentes.

Etiquetas de fila	Suma de Cantidad de casos
Caso - control	1
Cualitativos y cuantitativos.	13
Cuantitativa	1
Cuantitativos y cualitativos	1
Datos Cuantitativos	1
Mixto	1
S/D	2
Total general	20

Tabla 25. Metodologías de estudios empleados en las investigaciones.



Figura 4. Fecha de publicaciones de investigaciones por año.

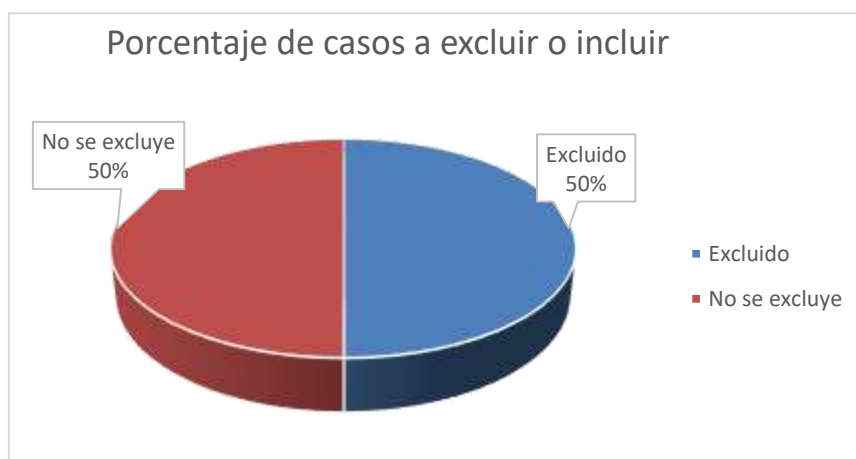


Figura 5. Porcentaje de casos que se evaluarán en la presente investigación.

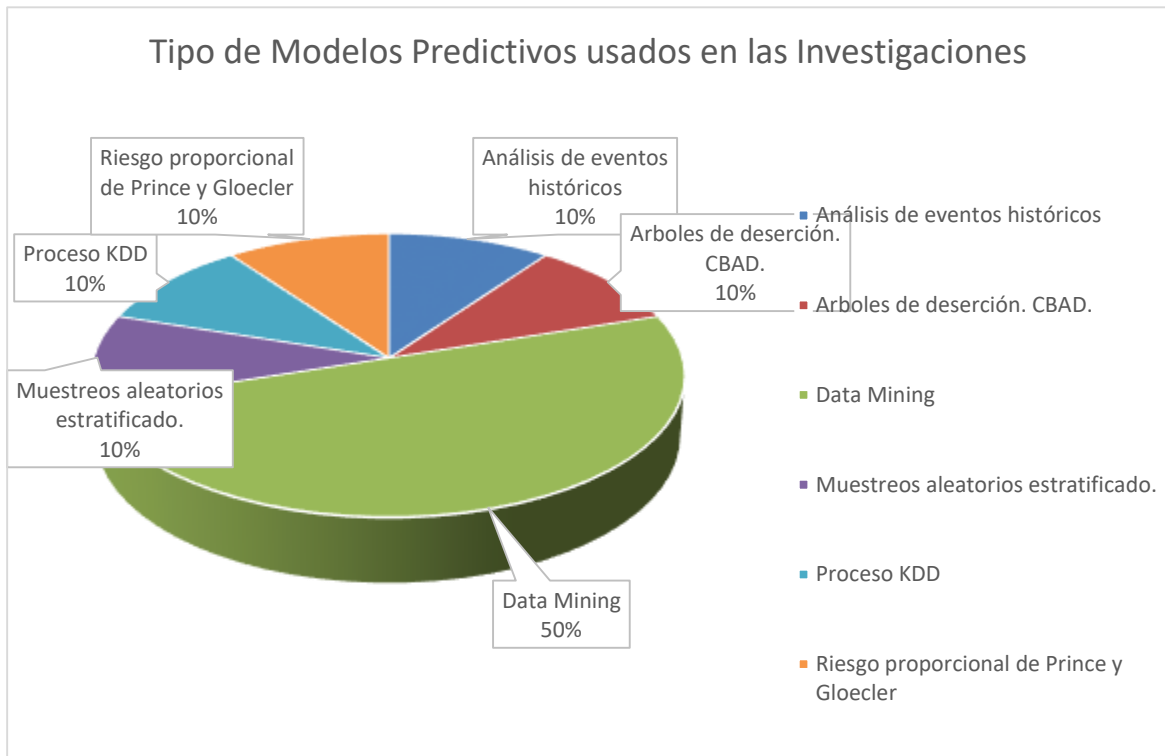


Figura 6. Tipos de Modelos predictivos usado en las Investigaciones.

Teniendo en cuenta la siguiente información de la *Figura 6* se detallará los tipos de modelos predictivos que son mencionados en esta.

Data Mining: es definido como un proceso de extracción de conocimientos útiles y comprensibles, que no es conocido, que se obtienen desde almacenes de datos muy grandes y de diferentes formatos.

Otra forma de poder describir el Data Mining como una explotación de búsquedas de información muy importante en grandes volúmenes de datos.

KDD: Knowledge Discovery in Databases, conocido como un proceso de descubrir información realmente útil y conocimientos dentro de una base de datos.

Muestreo aleatorio estratificado: técnica de muestreo que se puede utilizar cuando dentro de una población puedes distinguir claramente subgrupos o subpoblaciones. Mediante esta técnica de muestreo, la selección de los elementos que serán parte de la muestra serán separados dentro de un estrato.

Árboles de decisión: son posibles resultados que se pueden dar mediante una relación de decisiones. Se usa para poder encaminar ideas o anticipar matemáticamente mediante un algoritmo para una mejor opción.



Análisis de eventos históricos: Es la identificación de las variables que influyen en que ciertas situaciones se produzcan y esta, está controlado mediante una determinada unidad de tiempo.

CAPÍTULO IV. CONCLUSIONES

La revisión de literatura científica nos muestra que hoy en día los modelos predictivos son herramientas muy necesarias como apoyo a las entidades educativas para poder tener un conocimiento mayor de los motivos de deserción de los alumnos y de esta manera buscar una manera de mitigar dichas situaciones.

Se llega a la conclusión que el DATA MINING es una metodología preponderante durante el periodo 2008 y 2018 en todos los modelos de predicción como se puede visualizar en la *Figura 6*. Este modelo al juntarse con otras herramientas como árboles de decisión, CRISP-DM, así como apoyado con programación se puede hacer una herramienta mucho más fuerte y confiable.

Los métodos no harán que el alumno no se vuelva desertor, pero nos ayudará a tener conocimiento del riesgo que puede tener y cómo podemos ayudarlo para que no haga este cambio, por otro lado, cada persona (alumno) es un mundo distinto por lo que por más que tengamos herramientas cada uno tiene un poder de decisión según su entorno y otros factores.

Son muy pocas las universidades peruanas que cuentan con un modelo predictivo de deserción, es por ello que hemos podido tener una limitación de información, dado que sólo se ha podido encontrar un modelo predictivo a nivel nacional, a diferencia de nuestro vecino más cercano, Chile, que en la preocupación de sus estudiantes tanto de entidades públicas como privadas si cuentan con modelos de predicción que estimen la deserción de estudiantes.



REFERENCIAS

Mori-Sánchez MP. (2012) Deserción universitaria en estudiantes de una universidad privada de Iquitos. *Docencia Universitaria*, 6, 60-83.

Sánchez-Hernández, G., M.Barboza-Palomino, y H.Castilla-Cabello. (2017). Análisis de la deserción y los factores asociados a la permanencia estudiantil en una universidad peruana. *Actualidades Pedagógicas*, (69), 169-191.

Eckert, K., & Suénaga, R. (2015). Análisis de Deserción-Permanencia de Estudiantes Universitarios Utilizando Técnica de Clasificación en Minería de Datos. *Formación Universitaria*, 8 (5), 3-12.

Henríquez-Cabezas, N., & Escobar-Riffo, D. (2016). Construcción de un modelo de alerta temprana para la detección de estudiantes en riesgos de deserción de la universidad metropolitana de ciencias de la educación. *Revista Mexicana de Investigación Educativa*, 21 (71), 1221-1248.

DEL RÍO ARTEAGA, J. (2018). MODELO PREDICTIVO PARA LA RETENCIÓN DE ESTUDIANTES EN PRIMEROS AÑOS". TESIS PARA BACHILLER. UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA.

Miranda, M., & Guzmán, J. (2017). Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos. *Formación Universitaria*, 10 (3), 61-68.

Ramírez, Patricio E., & Grandón, Elizabeth E.. (2018). Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados. *Formación universitaria*, 11(3), 3-10.