

Modelos predictivos de la deserción estudiantil en una universidad privada peruana

OSWALDO SIFUENTES BITOCCHI¹

RECIBIDO: 11/04/2018 ACEPTADO: 20/08/2018

RESUMEN

La deserción es un problema que afecta a las universidades, públicas y privadas, y acarrea una serie de consecuencias negativas tanto para las instituciones como para los mismos jóvenes, por ello, el objetivo de este estudio fue determinar cómo el uso de modelos predictivos en asignaturas críticas contribuye a identificar a los estudiantes en riesgo de deserción. Se diseñaron siete modelos predictivos con la metodología CRISP (Cross-Industry Standard Process for Data Mining) y el historial académico de los estudiantes, para ser aplicados en siete cursos. Entre los principales resultados se puede destacar que los modelos predictivos contribuyeron a reducir en un 25 % y 40 % los niveles de desaprobación y las variables que mejor la predijeron fueron la carrera que estudian (vocación), el número de veces que se matriculan en la asignatura y la nota que tuvieron en matemática o comunicación cuando cursaron el quinto año de secundaria.

Palabras-claves: Deserción estudiantil; estudiantes universitarios; desaprobación; tutoría; modelos predictivos.

PREDICTIVE MODELS OF STUDENT DESERTION AT A PRIVATE PERUVIAN UNIVERSITY

ABSTRACT

Desertion is a problem that affects public and private universities, and leads to a series of negative consequences for both institutions and students. Therefore, the objective of this study was to determine how the use of predictive models in low pass-rate courses helps to identify students at risk of desertion. Seven predictive models were designed using CRISP (Cross-Industry Standard Process for Data Mining) methodology and students' academic records to be applied in seven low pass-rate courses. Among the main results, it can be noted that predictive models contributed to the reduction of fail rates by 25% and 40%, and that the variables that best forecast desertion were career choice (vocation), number of times students enrolled in the course, and grades obtained in mathematics or language arts when students attended the fifth year of high school.

Keywords: Student desertion; university students; fail; mentoring; predictive models.

1. INTRODUCCIÓN

La deserción es una situación en la que incurre un estudiante cuando deja de tener actividad académica durante tres semestres académicos consecutivos (Tinto, 1989), la cual genera diversos impactos negativos, tanto para al estudiante, porque repercute en sentimientos de frustración y fracaso que pueden limitar el desarrollo de sus competencias personales para la inserción laboral (Brock, 2010; Instituto Internacional para la Educación Superior en América Latina y el Caribe (IESALC), 2006), como para la universidad porque la disminución de ingresos provenientes de las matriculas pueden originar el deterioro de los estándares de calidad (Osorio, Bolancé y Castillo-Caicedo, 2012; Pascuas, Jaramillo y Verástegui, 2015); y a la sociedad porque la deserción contribuye a incrementar el desempleo y mantener los índices de pobreza.

La deserción puede darse en los primeros ciclos académicos o en los últimos, aunque la tendencia es mayoritaria en los periodos iniciales (Fontalvo, Castillo y Polo, 2014). Según Castaño, Gallón, Gómez y Vásquez (2004) y González (2005) se dan dos tipos de abandonos respecto al tiempo y al espacio. En el primer caso, la clasifican en: a) deserción precoz, cuando el estudiante es admitido por la universidad y no se matricula; b) deserción temprana, se da cuando el estudiante abandona sus estudios en los cuatro primeros semestres de la carrera; y c) deserción tardía, cuando el estudiante abandona los estudios a partir del quinto semestre en adelante.

Respecto al segundo, los investigadores dividen la deserción estudiantil en: a) deserción interna, ocurre cuando el estudiante decide cambiarse de carrera en la misma institución de educación superior; b) deserción institucional, se produce cuando el estudiante abandona la universidad para matricularse en otra; y c) deserción definitiva del sistema educativo, se da cuando el estudiante no se matricula en ninguna universidad.

Sobre las causas de la deserción universitaria, algunos estudiosos sostienen que estas pueden darse por razones económicas, sociales e institucionales, mientras que otros agregan a los factores psicológicos, sociológicos y de interacciones (Himmel, 2012). Cabrera, Bethencourt, Álvarez y González (2006) las

¹ Magíster en Administración de Negocios, Director de Gestión Académica de la Universidad Continental
E-mail: osifuentes@continental.edu.pe

agrupan en cinco categorías: a) psicológico, referida a los rasgos de personalidad, creencias y actitudes propios del estudiante para la culminación de sus estudios; b) sociológico, la falta de integración de los estudiantes en el entorno de la educación superior puede llevarlo a desertar (Spady, 1971); c) económico, los beneficios sociales y económicos que pueden producir los estudios universitarios se perciben como inferiores a los de otras actividades (Donoso y Schiefelbein, 2007); d) organizacional, involucra las características de la universidad, la calidad docente y las experiencias en las aulas (Donoso y Schiefelbein, 2007); e) interaccionista, a mayor interacción entre estudiantes y profesores, más probabilidades que los estudiantes finalicen sus estudios (Tinto, 1987).

Asimismo, esta problemática trae serias consecuencias en el aspecto socioeconómico del país, ya que se continúa con el círculo de la pobreza y el desempleo (Osorio et al., 2012), porque lo más probable es que los estudiantes desertores tengan pocas posibilidades de inclusión y de mejorar su calidad de vida; pero a pesar de ello son pocos los estudios realizados sobre la deserción, que puedan contribuir a tener una estadística exacta en el Perú, salvo estudios como el de (Mori, 2013), quien manifiesta que la tasa anual de ausentismo estudiantil universitario alcanza el 17 %; y estudios específicos en las carreras profesionales de salud que señalan un 10 % (Heredia et al., 2015).

Como alternativa o estrategia para identificar los principales factores que determinan la deserción estudiantil, algunos investigadores han diseñado modelos matemáticos o predictivos, que son definidos como instrumentos que se elaboran en base a la información relevante o datos experimentales de las variables que se están investigando, y con el análisis de dichas variables predictoras ayudan a predecir resultados para condiciones no estudiadas (Fernández *et al.*, 1997, citado en Santiesteban-López y López-Malo, 2008).

Para la elaboración de dichos modelos, estudiosos como Eckert y Suénaga (2015), Azoumana (2013) y Rodallegas *et al.* (2010) trabajaron con el historial académico (asignaturas aprobadas, cantidad y resultado de asignaturas cursadas, procedencia y edad de ingreso del estudiante, etc.), apoyados con la técnica de minería de datos, la cual permite de manera dinámica analizar grandes cantidades de datos para obtener patrones de comportamiento o tendencias ocultas que pueden ser muy útiles para la toma de decisiones (Escribano-García, Martínez-de-Pisón, Castejón-Limas, Sanz-García y Fernández-Martínez, 2010).

En este artículo se presentan los resultados obtenidos de los diseños y aplicaciones de siete modelos predictivos correspondientes al mismo número de cursos, denominados por la universidad privada como críticos por alto índice de reprobación, que es una de las causas de la deserción, con el objetivo de determinar cómo sus usos contribuyen a identificar a los estudiantes en riesgo de deserción, al inicio del ciclo académico, para que la institución desarrolle estrategias de apoyo sobre todo para los estudiantes con mayores probabilidades de desertar.

2. METODOLOGÍA

Este estudio es descriptivo de tipo selectivo de corte transversal (Ato, López y Benavente, 2013). Con la información histórica académica de los 4478 estudiantes matriculados en el ciclo 2016-I (2731 varones y 1747 mujeres, con una edad media de 23 años), de diversas carreras profesionales de una universidad privada peruana, se elaboraron siete modelos predictivos para igual número de cursos, considerados como críticos por los altos índices de reprobación.

Diseño de modelos predictivos

Con la técnica de minería de datos se revisaron los patrones de rendimiento académico de los estudiantes, tanto de los aprobados como desaprobados, para poder determinar los cursos en los que se trabajarían los modelos predictivos. De todos los cursos con mayor índice de desaprobados, se seleccionaron de manera aleatoria siete de ellos: Estadística II, Introducción a la comunicación, Introducción a la matemática, Matemática I, Pre-cálculo, Química y Resistencia de materiales.

Posteriormente se utilizó la metodología CRISP (Cross-Industry Standard Process for Data Mining), la cual incluye seis etapas como proceso de modelamiento para determinar una selección de las mejores variables predictoras del buen rendimiento académico. Seguidamente, se aplicó el método de crecimiento árboles de clasificación y regresión (Classification and Regression Trees), para dividir los datos en segmentos con la finalidad de que sean lo más homogéneos posibles respecto a la deserción para el análisis de clasificación exploratorios y confirmatorios. Entre las variables identificadas con mayor recurrencia están: género, edad, tipo de colegio, promedio académico ponderado, departamento del colegio de procedencia, ingreso familiar mensual y el monto de la pensión que pagan los estudiantes en la universidad privada (tabla 1).

Tabla 1. Variables consideradas para los modelos predictivos de los siete cursos

Variable independiente	Importancia
Promedio ponderado acumulado (antes del semestre)	100.0 %
Ratio (promedio ponderado acumulado aprobados)	84.4 %
Promedio ponderado del semestre anterior	59.2 %
Carrera que estudia	17.2 %
Número de veces de matrícula en el curso	13.1 %
Turno	11.1 %
Edad	6.4 %
Sexo	6.3 %
Monto de la última pensión	4.5 %
Créditos matriculados (en el semestre)	3.4 %

Posteriormente, al comenzar el semestre académico 2016-II se ingresó la información de los nuevos estudiantes que se matricularon en los siete cursos mencionados anteriormente y se hizo la corrida de los modelos predictivos, con los porcentajes de deserción obtenidos por cada uno de los cursos críticos la institución de educación superior que participa en el estudio ejecutó estrategias de acompañamiento para apoyar a los estudiantes con mayor tendencia a desertar.

3. RESULTADOS

De las variables que se consideraron en el diseño de los siete modelos predictivos: edad, sexo, colegio de procedencia, departamento del colegio de procedencia, ingreso familiar mensual, número de veces de matrícula en la asignatura, promedio ponderado y monto de la pensión que pagan los estudiantes en la universidad, las que mejor predijeron la deserción fueron: la carrera que estudian los estudiantes, es decir, la elección de la carrera, el número de veces que se matriculan en la asignatura y la nota que obtuvieron en los cursos de matemática o comunicación cuando cursaron el quinto año del nivel secundaria.

Los resultados obtenidos con la corrida de los siete modelos predictivos, diseñados para la presente investigación, predijo que 2777 estudiantes desaprobarían los siete cursos considerados en el estudio, de un total de 4478 jóvenes matriculados en las mencionadas asignaturas en ciclo académico 2016-II. Con dicha información la universidad, a través de sus oficinas de acompañamiento al estudiante, brindó apoyo durante todo el semestre académico al grupo de jóvenes con mayores probabilidades de reprobación y al finalizar el semestre lograron que 1210 estudiantes de ese grupo aprobaran y solo desaprobaron 1567 (tabla 2).

Las estadísticas de desaprobación de los últimos años muestran que se ha producido un ligero incremento en ese indicador, pero que a partir de la implementación de los modelos predictivos creados para los siete cursos calificados como críticos, así como la aplicación de estrategias de apoyo al estudiante, se ha reportado una disminución en el nivel de desaprobación, estimado como uno de los factores de la deserción. Las estadísticas de los siete cursos se muestran en la tabla 3, considerando que los modelos se implementaron en el periodo 2016-II.

Se desprende de la tabla, por ejemplo, que si en el segundo semestre del 2014 en el curso de Matemática I reprobaban 1661 estudiantes y solo aprobaban 861 de ellos, dos años después, con la corrida de los modelos predictivos y la implementación de acciones de acompañamiento a los estudiantes, en la misma asignatura desaprobaron 233 y aprobaron 427, invirtiéndose las cifras en comparación con el primer año mencionado. En la misma relación, en

Tabla 2. Estimación de los modelos predictivos y resultados alcanzados

Cursos	Predicho por modelo		Resultados al finalizar el ciclo	
	Probables aprobados	Probables desaprobados	Aprobados	Desaprobados
Estadística II	210	831	652	389
Introducción a la comunicación	171	82	167	86
Introducción a la matemática	92	214	175	131
Matemática I	270	390	427	233
Pre-cálculo I	191	625	485	331
Química	424	357	549	232
Resistencia de materiales	343	278	456	165
Total	1701	2777	2911	1567

Tabla 3. Índices de desaprobación antes y después de los modelos predictivos

Cursos	2014-II		2015-II		2016-II	
	Aprobados	Desaprobados	Aprobados	Desaprobados	Aprobados	Desaprobados
Estadística II	227	744	539	968	652	389
Introducción a la comunicación	81	92	215	217	167	86
Introducción a la matemática	78	93	360	378	175	131
Matemática I	861	1661	723	722	427	233
Pre-cálculo I			595	793	485	331
Química	83	32	820	767	549	232
Resistencia de materiales	137	101	143	80	456	165
Total	1467	2723	3395	3925	2911	1567

Fuente: Oficina de registros académicos de la universidad privada que participa en el estudio.

Nota: El curso de Pre-cálculo no se dictaba en el ciclo 2014-II.

Tabla 4. Índices de desaprobación con los modelos predictivos en %

Cursos	2014-II		2015-II		2016-II	
	Aprobados	Desaprobados	Aprobados	Desaprobados	Aprobados	Desaprobados
Estadística II	23%	77%	36%	64%	63%	37%
Introducción a la comunicación	47%	53%	50%	50%	63%	37%
Introducción a la matemática	46%	54%	49%	51%	55%	45%
Matemática I	34%	66%	50%	50%	59%	41%
Pre-cálculo I			43%	57%	56%	44%
Química	72%	28%	52%	48%	69%	31%
Resistencia de materiales	58%	42%	64%	36%	55%	45%
Total	35%	65%	46%	54%	60%	40%

el curso de Estadística II el margen de diferencia entre aprobados y desaprobados se incrementó a favor de los aprobados en un 40% (tabla 4).

4. DISCUSIÓN

Los resultados obtenidos demuestran que los modelos predictivos ayudan a predecir de manera temprana, según las características e historial académico de los estudiantes, cuántos jóvenes universitarios tienen mayores probabilidades de reprobado una asignatura y que el apoyo que les pueda brindar la universidad es fundamental para que ellos logren aprobar los cursos. Lo mencionado concuerda con los resultados obtenidos por Eckert y Suénaga (2015), quienes hallaron en su investigación que el apoyo que les pueda brindar la institución es crucial durante el primer año y que la cantidad de asignaturas aprobadas marca una tendencia a lo largo de la carrera.

Para Herrero, Merlino, Ayllón y Escanés (2013), también es importante que los estudiantes aprueben la mayor cantidad de cursos posibles durante

los primeros ciclos académicos porque eso permite reducir el riesgo de deserción en los siguientes semestres. Los investigadores trabajaron con modelos de duración en programas de prevención al inicio del ciclo, como es el caso de la presente investigación, con la finalidad de reducir los índices de deserción al finalizar dicho periodo de estudio.

Sin duda, la aprobación de los cursos contribuye significativamente en la permanencia estudiantil, por ende, se produce un menor porcentaje de deserción, que en los últimos años se ha convertido en un serio problema para las instituciones de educación superior, tanto públicas como privadas, lo que también acarrea un problema social y económico para el país, ya que los estudiantes desertores tendrán menos posibilidades de conseguir un trabajo, mejor remunerado, que les permita sostener a sus familias y mejorar sus condiciones de vida.

Los hallazgos encontrados respecto a la elección de la carrera o vocación, como una variable determinante en la deserción, coinciden con las investigaciones de Heredia *et al.* (2015) y Herrero *et al.* (2013), quienes, conjuntamente con Celis Schnei-

der *et al.* (2013), también consideran al factor económico (ingreso mensual de los padres y pensión que pagan los estudiantes) como un aspecto influyente en la deserción, mientras que en el presente estudio esa variable no resulta crucial, aunque sí tiene una moderada asociación con el abandono de los estudios universitarios.

En cuanto a la variable del sexo, si el ser varón o mujer influye en la deserción, investigaciones como las de Osorio *et al.* (2012) afirman que sí tiene un alto nivel de implicancia en la deserción, pero en este estudio se halló que el factor mencionado tiene un bajo nivel de predicción en el abandono estudiantil, así como la edad de los estudiantes, aspecto que Rodallegas *et al.* (2010) consideró en la construcción de su modelo y resultó ser determinante para predecir la reprobación de los cursos.

5. CONCLUSIONES

Si bien es cierto la deserción estudiantil es un problema latente en las diversas universidades peruanas, públicas y privadas, pero existen mecanismos o estrategias que pueden contribuir en contrarrestarla, como se ha demostrado en la presente investigación, que con la elaboración y diseño de modelos predictivos, basado en las características propias de los estudiantes de cada institución, se puede detectar de manera temprana a los jóvenes que tienen mayores riesgos de reprobar algún curso y que luego podría llevarlo a desertar de la vida universitaria.

Con esta predicción, una vez realizada la corrida de los siete modelos predictivos al inicio del semestre académico, la institución de educación superior puede tomar las acciones que considere pertinentes para ayudar al estudiante, en este caso la universidad privada en estudio realizó un acompañamiento tutorial durante las 16 semanas del ciclo académico, esfuerzos que fueron focalizados en los estudiantes que tenían poca probabilidad de aprobar las asignaturas.

Es una labor que trabajaron coordinadamente con las diversas áreas que prestan apoyo a los jóvenes estudiantes como la oficina de tutoría y orientación psicológica conjuntamente con la dirección académica de la carrera y la dirección de gestión docente, esfuerzo que se vio recompensado al finalizar el semestre porque lograron reducir la desaprobación de los cursos que participaron en el estudio en un 25% a 40%, en comparación con los años anteriores en los que no se trabajaron con modelos predictivos.

La efectividad de los siete modelos predictivos se logró porque se trabajó cada modelo específicamente para cada curso, debido a que las características de las asignaturas y las necesidades de los estudiantes eran distintas, factores que fueron analizados en el historial académico de los estudiantes proporcionado por la oficina de registros de matrículas de la universidad privada que participó en el estudio.

6. REFERENCIAS BIBLIOGRÁFICAS

- [1] Ato, M., López, J. J., y Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059.
- [2] Azoumana, K. (2013). Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos. *Pensamiento Americano*, 6(10), 41-51.
- [3] Brock, T. (2010). Young Adults and Higher Education: Barriers and Breakthroughs to Success. *The Future of Children*, 20(1), 109-132.
- [4] Cabrera, L., Bethencourt, J., Álvarez, P., y González, M. (2006). El problema del abandono de los estudios universitarios. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 12(2), 171-203.
- [5] Castaño, E., Gallón, S., Gómez, K., y Vásquez, J. (2004). Deserción estudiantil universitaria: una aplicación de modelos de duración. *Lecturas de Economía*, (60), 39-65.
- [6] Celis Schneider, R., Ramírez, F., Luz, C., Martínez, R., Cristina, M., y Venegas Villanueva, H. (2013). Factores de riesgo de deserción presentes en alumnos repitentes de las carreras de enfermería y kinesiología en una universidad chilena. *Ciencia y enfermería*, 19(3), 63-71.
- [7] Donoso, S., y Schiefelbein, E. (2007). Análisis de los modelos explicativos de retención de estudiantes en la universidad: Una visión desde la desigualdad social. *Estudios pedagógicos (Valdivia)*, 33(1), 7-27.
- [8] Eckert, K. B., y Suénaga, R. (2015). Análisis de deserción-permanencia de estudiantes universitarios utilizando técnica de clasificación en minería de datos. *Formación universitaria*, 8(5), 03-12.

- [9] Escribano-García, R., Martínez-de-Pisón, F., Castejón-Limas, M., Sanz-García, A., y Fernández-Martínez, R. (2010). Modelos descriptivos y predictivos para la estimación de costes en proyectos informáticos. Conferencia llevada a cabo en *XIV International Congress on Project Engineering* (pp. 2590-2600), Madrid, España.
- [10] Fontalvo, W., Castillo, M., y Polo, S. (2014). Análisis comparativo entre las características más relevantes de deserción estudiantil en el programa de Ingeniería Industrial de la Universidad Autónoma del Caribe: Estudiantes activos en el periodo 2013-01 y desertores académicos de los periodos 2011-01 a 2012-02. *Escenarios*, 12(1), 96-104.
- [11] González, L. (2005). *Estudio sobre la repitencia y deserción en la educación superior chilena*. Digital Observatory for higher education in Latin America and The Caribbean. IESALC-UNESCO. Recuperado de <http://unesdoc.unesco.org/images/0014/001400/140087s.pdf>
- [12] Heredia, M., Andía, M., Ocampo, H., Ramos-Castillo, J., Rodríguez, A., Tenorio, C., y Pardo, K. (2015). Deserción estudiantil en las carreras de ciencias de la salud en el Perú. *Anales de la Facultad de Medicina*, 76(SPE), 57-61.
- [13] Herrero, V., Merlino, A., Ayllón, S., y Escanés, G. (2013). Aplicación de un modelo de duración en programas de prevención de deserción universitaria. *Revista Electrónica de Investigación Educativa*, 15(3), 38-52.
- [14] Himmel, E. (2012). Modelos de análisis de la deserción estudiantil en la educación superior. *Revista Calidad de la Educación*, (17), 91-108.
- [15] Instituto Internacional para la Educación Superior en América Latina y el Caribe (IESALC). (2006). *Informe sobre la educación superior en América Latina y el Caribe 2000-2005: La metamorfosis de la educación superior* (Institucional) (p. 262). Caracas: Instituto Internacional para la Educación Superior en América Latina y el Caribe (IESALC). Recuperado de <http://C:/Users/mir/Downloads/InformeES-2000-2005.pdf>
- [16] Mori, M. (2013). Deserción universitaria en estudiantes de una universidad privada de Iquitos. *Revista Digital de Investigación en Docencia Universitaria*, 6(1), 60-83.
- [17] Osorio, A., Bolancé, C., y Castillo-Caicedo, M. (2012). Deserción y graduación estudiantil universitaria: una aplicación de los modelos de supervivencia. *Universia*, 3(6), 31-57.
- [18] Pascuas, Y. S., Jaramillo, C. O., y Verástegui, F. A. (2015). Desarrollo de objetos virtuales de aprendizaje como estrategia para fomentar la permanencia estudiantil en la educación superior. *Revista Escuela de Administración de Negocios*, (79), 116-129.
- [19] Rodallegas, E., Torres, A., Gaona, B., Gastelloú, E., Lezama, R., & Valero, S. (2010). Modelo predictivo para la determinación de causas de reprobación mediante Minería de Datos. *Recursos digitales para la educación y la cultura, Kaambal*, 48-55.
- [20] Santiesteban-López, N., y López-Malo, A. (2008). Descripción e importancia de algunos modelos predictivos utilizados como herramienta para la conservación de alimentos. *Temas Selectos de Ingeniería de Alimentos*, 2(2), 14-26.
- [21] Spady, W. (1971). Dropouts from Higher Education: Toward an Empirical Model. *Interchange*, 2(3), 38-62.
- [22] Tinto, V. (1987). The Principles of Effective Retention. Presentado en Fall Conference of the Maryland College Personnel Association, Prince George's Community College, Largo, MD.: ERLC. Recuperado de <https://files.eric.ed.gov/fulltext/ED301267.pdf>
- [23] Tinto, V. (1989). Definir la Deserción: Una Cuestión de Perspectiva. *Revista de la educación superior*, (71), 33-51.