

[Cierre de edición el 01 de Enero del 2019]

doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

## Un análisis multinomial y predictivo de los factores asociados a la deserción universitaria

*A Multinomial and Predictive Analysis of Factors Associated with University Dropout*

*Uma análise multinomial e preditiva dos fatores associados a desistências universitárias*

*Tatiana Fernández-Martín*

Instituto Tecnológico de Costa Rica

Cartago, Costa Rica

[tafema@itcr.ac.cr](mailto:tafema@itcr.ac.cr)

 <https://orcid.org/0000-0002-7726-0068>

*Martín Solís-Salazar*

Instituto Tecnológico de Costa Rica

Cartago, Costa Rica

[marsolis@itcr.ac.cr](mailto:marsolis@itcr.ac.cr)

 <https://orcid.org/0000-0003-4750-1198>

*María Teresa Hernández-Jiménez*

Instituto Tecnológico de Costa Rica

Cartago, Costa Rica

[thernandez@itcr.ac.cr](mailto:thernandez@itcr.ac.cr)

 <https://orcid.org/0000-0002-5672-5743>

*Tania Elena Moreira-Mora*

Instituto Tecnológico de Costa Rica

Cartago, Costa Rica

[tmoreira@itcr.ac.cr](mailto:tmoreira@itcr.ac.cr)

 <http://orcid.org/000-0002-8955-0804>



Recibido • Received • Recebido: 25 / 01 / 2017

Corregido • Revised • Revisado: 07 / 07 / 2018

Aceptado • Accepted • Aprovado: 20 / 09 / 2018



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

**Resumen:** El fenómeno de la deserción por su complejidad e impacto educativo y social ha sido ampliamente estudiado para entender las causas específicas. Dentro de esta línea de investigación, el objetivo de este estudio fue desarrollar modelos explicativos y predictivos de la deserción estudiantil en los programas académicos de grado del Instituto Tecnológico de Costa Rica, con base en diferentes variables registradas en el Sistema de Indicadores de Gestión Institucional (SIGI). En una primera etapa se procesan modelos de regresión multinomial para identificar la influencia de estas variables en la deserción. Mientras que en la segunda se evaluaron seis algoritmos de aprendizaje automático con el fin de buscar un modelo que permitiera detectar al estudiantado desertor. Entre los hallazgos se destaca que la probabilidad de deserción se relaciona con variables sociodemográficas, programa de estudio, beneficios obtenidos al ingresar, historial académico y rendimiento en el primer semestre de estudio. Además, resultó que el mejor algoritmo predictor de la deserción fue el “random forest”, por identificar correctamente al potencial desertor con una probabilidad de 0.83 y por capturar al 34% de deserción real. Estos resultados son el primer paso hacia la construcción de un modelo predictivo más robusto de la deserción estudiantil, que se espera contribuya a la toma de decisiones preventivas en esta universidad.

**Palabras claves:** Multinomial; modelos predictivos; educación superior; deserción estudiantil; factores institucionales y estudiantiles.

**Abstract:** The phenomenon of dropout, by its complexity and educational and social impact, has been extensively studied to understand the specific causes. In this line of research, the purpose of this study was to analyze explanatory and predictive models of student dropout from university studies at the Instituto Tecnológico de Costa Rica (TEC), based on many variables recorded in the institutional system indicators. The first stage of the analysis considered multinomial regression models to identify the influence of these variables on the dropout. In the second analysis, six machine learning algorithms were evaluated in order to find a model that would predict student dropout. Data analysis showed that the probability of dropping out is related to sociodemographic variables, study program, academic history, scholarship and other benefits, and performance after first semester. In addition, the best predictor of dropout algorithm was the “random forest”, a probability of 0.83 to predict the dropout correctly and to capture 34% of the actual student dropout. These results are the first step toward building a more robust predictive model of dropout, which will contribute to preventive decision making in this university.

**Keywords:** Multinomial; student dropout; predictor models; higher education; institutional and student’s factors.

**Resumo:** O fenômeno da deserção, por sua complexidade e impacto educacional e social tem sido amplamente estudado a fim de compreender suas causas específicas. Dentro desta linha de pesquisa, o objetivo deste estudo foi desenvolver modelos explicativos e preditivos de deserção estudiantil nos cursos de graduação do Instituto Tecnológico da Costa Rica, com base em diferentes variáveis registradas no Sistema de Indicadores de Gestão Institucional (SIGI). No primeiro estágio, os modelos de regressão multinomial são processados para identificar a influência dessas variáveis na deserção. Enquanto no segundo, seis algoritmos de aprendizado de máquina foram avaliados, a fim de encontrar um modelo que permitisse um aluno ser identificado como um desertor. Como resultado,

doi: <http://dx.doi.org/10.15359/ree.23-1.5>URL: <http://www.una.ac.cr/educare>CORREO: [educare@una.cr](mailto:educare@una.cr)

destaca-se que a probabilidade de deserção está relacionada às variáveis sociodemográficas, programa de estudos, benefícios obtidos na hora do ingresso, registro acadêmico e desempenho no primeiro semestre de estudo. Além disso, descobriu-se que o melhor preditor de deserção foi o “random forest”, por identificar corretamente o potencial desertor com uma probabilidade de 0,83 e por capturar 34% da deserção real. Esses resultados são o primeiro passo para a construção de um modelo preditivo mais eficaz de deserção estudantil, que poderia contribuir para a tomada de decisões preventivas nessa universidade.

**Palavras-chave:** Multinomial; modelos preditivos; educação superior; deserção de estudantes; fatores institucionais e estudantis.

## Introducción

En el ámbito de la educación superior estatal, una tarea obligatoria es el análisis del fenómeno de la deserción “no solo para determinar las causas específicas ... sino también determinar en qué medida la institución puede tomar acciones preventivas; [para reducir la tasa de abandono]; pues es probable ... [que] no sea solo un problema personal del estudiante” (Castillo, 2010, p. 54). A pesar de que esta temática no es novedosa o ajena para quienes conocen el entorno universitario; según Castaño, Gallón, Gómez y Vásquez (2006, p. 10) “las altas tasas de deserción y bajas tasas de graduación se han convertido en un asunto de creciente interés para las instituciones de educación superior y las autoridades educativas” en los últimos cincuenta años, dado que este fenómeno amplía la brecha social y económica, así como limita el desarrollo de la nación.

A este respecto, el estudio del Centro Interuniversitario de Desarrollo (CINDA, 2006) expone un índice de deserción universitaria general cercano al 57% para América Latina. Por su parte, los índices más altos de deserción:

... en América Latina, según Botero (2006) los datos de deserción en Guatemala alcanzan el 82%, en Bolivia el 73,3%, en Uruguay el 72%, en Brasil el 69%, en Costa Rica el 54%, en Chile el 53,7%, en México el 53%, en Panamá el 42%, en Argentina el 40% y en Honduras el 38%, de otra parte un estudio de la UNESCO/IESALC (2006) mostró que el índice de deserción en América Latina bordea el 57%; el país donde la tasa de deserción es menor es Cuba, con un 25%. (Villamizar y Pérez, 2011, p. 133)

Estos datos no distan de las tasas de abandono en países como España, Estados Unidos, Francia y Austria, ubicados entre un 30% y un 50%; sin embargo, sí se presenta una disminución importante de estos porcentajes en países como Alemania con un 25%, los Países Bajos con un 20% y Suiza o Finlandia que presentan un índice entre 7% y 10% de deserción (Vélez y López, 2004).



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

En general, sobresalen las investigaciones descriptivas y correlacionales de las variables asociadas al fenómeno de la deserción, como los estudios de [Brenes \(2005\)](#), [Merlino, Ayllón y Escanés \(2011\)](#), [Regueyra y Rojas \(2013\)](#) y [Treviño, Ibarra, Castán, Laria y Guzmán \(2013\)](#), quienes coinciden en que la deserción estudiantil debe ser explicada a partir del análisis de múltiples factores: personales, familiares, económicos, institucionales y los referidos a los temas vocacionales. En la educación estatal de Costa Rica se ha encontrado que:

Poco menos de la mitad de los estudiantes que ingresan a una institución universitaria estatal logra obtener un grado académico dentro de ese mismo sector, en comparación con alrededor de un 70% de las dos universidades privadas que se estudiaron. Además, solamente el 10% de quienes se gradúan lo hacen en el tiempo esperado según los planes de estudio de las distintas carreras. ([Brenes, 2005, p. 2](#))

[Abarca y Sánchez \(2005\)](#) realizan un perfil del estudiantado desertor de la Universidad de Costa Rica de las cohortes 1993 a 1998, reportan porcentajes de deserción entre el 45% y el 36%, y ubican como áreas de mayor deserción las ciencias básicas, artes y letras y en tercer lugar las ingenierías. Además, concluyen que el no ingresar a la carrera preferida es uno de los mayores alicientes para el abandono; sin embargo, rescatan que a pesar de abandonar la universidad en la que se realiza el estudio, el estudiantado, en su mayoría, no deserta del sistema universitario. Por su parte, [Regueyra y Rojas \(2013\)](#) realizan una investigación en ese mismo centro de estudios para la Escuela de Trabajo Social y encontraron como principal causa el haber ingresado a una carrera que no era de su interés.

La [Universidad Nacional \(2008\)](#) publica también un estudio respecto al tema de deserción, tomando como estudiantado desertor al de “primer ingreso” que se incorporó a la institución en los años 2003, 2004, 2005 y 2006 por el proceso de admisión y que en los años siguientes a su ingreso no matriculó ni obtuvo ningún grado. En este mismo se identifican como principales causas de la deserción la falta de gusto por la carrera matriculada (42%), la falta de recursos económicos (23%), los horarios en los que se ofrecen los cursos (17%) y razones personales-familiares (16%), entre otras.

Específicamente en el Instituto Tecnológico de Costa Rica (TEC), los estudios y datos respecto a la deserción estudiantil surgen desde los primeros años de la institución. Para 1978, 1979 y 1980 el [Instituto Tecnológico de Costa Rica, División de Servicios Estudiantiles y Académicos, Departamento de Orientación \(1980a, 1980b, 1981\)](#) proporciona los primeros informes que documentan la deserción, en los cuales identifican como posibles causas esencialmente las referidas a insatisfacción con la carrera, traslados a otras universidades, principalmente a la Universidad de Costa Rica y problemas económicos. Por otra parte, [Barquero \(1997\)](#) presenta datos de la deserción acumulada de la institución para diversas cohortes; el primer caso es de la cohorte de 1981 que al II semestre de 1988 presentaba alrededor de un

doi: <http://dx.doi.org/10.15359/ree.23-1.5>URL: <http://www.una.ac.cr/educare>CORREO: [educare@una.cr](mailto:educare@una.cr)

64% de deserción, mientras que la cohorte 1987 para el II semestre de 1996 presentaba un 63% de estudiantes que habían desertado. Además, [Barquero \(1997\)](#) hace un recuento de las principales causas de deserción, a saber: problemas laborales (49%), problemas económicos (22%), problemas familiares (16%) y traslado a otra universidad (13%), entre otras. Un elemento importante identificado en estos estudios es que alrededor del 50% del estudiantado que se retiraba de la institución lo hacía en los primeros dos años de carrera.

Más recientemente, [Chinchilla \(2013\)](#) realiza un estudio sobre deserción con la cohorte 2012, incluyendo estudiantes que no realizaron su matrícula para el II semestre del mismo año o al I semestre del año 2013; dentro de los resultados se destaca que un 52% de la población consultada desertó en el primer semestre y las razones principales de deserción tienen que ver con no estar en la carrera deseada (37,2%), seguido por el traslado a otra universidad (15,9%) y razones económicas (10,6%), entre otras.

En suma, a pesar de las diferencias en cada uno de los contextos citados hay coincidencias en sus hallazgos. En la categoría estudiantil se han encontrado evidencias de la relación entre la deserción y los antecedentes académicos, los asuntos económicos, personales o familiares y los aspectos referidos al área psicoeducativa -como la motivación, organización, autodisciplina, responsabilidad, autoeficacia y satisfacción con la universidad-. Mientras que en la categoría institucional se han encontrado evidencias con el ambiente de la clase y aspectos docentes referidos a la evaluación y estrategias didácticas o pedagógicas, la sobrecarga en el plan de estudios o dificultad de la carrera. Igualmente, con el apoyo brindado por la institución en el financiamiento, apoyo estudiantil y acciones preventivas que desarrollan las instituciones o aspectos administrativos. Finalmente, factores externos como crisis económicas mundiales o nacionales y eventos políticos deben ser considerados en estos análisis, como se desprende de la revisión de datos capturados en el Sistema Institucional de Gestión de Indicadores (SIGI) por el [TEC, Oficina de Planificación Institucional \(2014\)](#).

Precisamente este estudio tiene como objetivo desarrollar modelos explicativos y predictivos de la deserción estudiantil en los programas académicos de grado del Tecnológico de Costa Rica, con base en diferentes variables registradas en el SIGI, considerando tanto la población desertora como activa de tres cohortes universitarias. Los modelos explicativos pretenden comprender la relación existente entre la deserción y un conjunto de variables relacionadas con: características sociodemográficas del estudiantado, del programa de estudio, beneficios obtenidos por el estudiantado al ingresar, historial académico, e historial del primer ciclo lectivo. Por su parte, los modelos predictivos pretenden identificar un algoritmo que ayude a predecir anticipadamente si el estudiante va a desertar. En general, estos modelos permiten identificar patrones y asociaciones entre las variables incluidas en los algoritmos y la deserción ([Pal, 2012](#)). En el caso de esta investigación, estos modelos fueron estimados con las mismas variables utilizadas en el modelo multinomial. La mayor parte de las investigaciones sobre deserción se enfocan en intentar explicarla o predecirla, pero no abordan ambas problemáticas.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

El comprender el fenómeno y, a la vez, predecirlo permitirá responder al problema de esta investigación enfocado en identificar cuáles son los factores asociados a la deserción y cómo detectar al futuro estudiantado desertor. Ambos focos de conocimiento se vuelven relevantes para intervenir y prevenir esta problemática en la educación superior. Según se estipula en el eje de cobertura y equidad de PLANES 2016-2020 (CONARE, 2016), las universidades estatales deben cumplir un conjunto de políticas y acciones que aseguren el acceso, la permanencia y el éxito académico de la población estudiantil. En esta línea, la presente investigación responde al planteamiento de la política específica institucional 1.1.1 (TEC, 2014), enfocada en garantizar la equidad de oportunidades y permanencia exitosa en la oferta académica.

En el caso del Instituto Tecnológico de Costa Rica, esta problemática se muestra en los datos de la [Tabla 1](#), donde se contrasta el total de estudiantado matriculado (MT) de primer ingreso con no matriculado (NO MT) en el segundo semestre para los años comprendidos entre 2010 y 2014.

**Tabla 1:** Porcentaje de estudiantes de primer ingreso del ITCR por año, que no matricularon en el segundo semestre del mismo año

Grado académico	Años														
	2010			2011			2012			2013			2014		
	MT	NO MT	%												
Bachillerato	579	69	11,9	435	80	18,4	453	70	15,4	722	107	14,8	641	117	18,3
Licenciatura	809	96	11,9	651	116	17,8	848	120	14,2	1.183	113	9,6	1.049	137	13,1
TOTAL	1.388	165	11,9	1.086	196	18,1	1.301	190	14,6	1.905	220	11,6	1.690	254	15,0

**Nota:** Sistema de Indicadores de Gestión Institucional. Elaborado con información de Bases de Datos del Departamento de Admisión y Registro, generado al 24 de febrero de 2015 (Instituto Tecnológico de Costa Rica, Oficina de Planificación Institucional, 2018).

Como lo muestran estos datos, el porcentaje de estudiantes de primer ingreso en carreras de grado que no se matricularon en el segundo semestre osciló entre el 11,55% y 18,05% durante este periodo. Estos datos iniciales muestran la relevancia de profundizar en esta problemática y en sus posibles factores a partir de modelos explicativos y predictivos.

## Método

El estudio del fenómeno de la deserción en el TEC se realizó desde un enfoque metodológico cuantitativo basado en un diseño correlacional predictivo, específicamente con la estimación de un modelo explicativo y predictivo exploratorio a partir de las siguientes variables individuales e institucionales.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

- Deserción: De acuerdo con el [Consejo Nacional de Rectores \(CONARE, 2008\)](#) de Costa Rica, un desertor o desertora es aquella persona que no matricula ningún curso en la universidad por un periodo consecutivo de dos años. Con base en esta definición, se clasifican los sujetos de estudio en tres categorías:
  - Estudiantado no desertor: incluye estudiantes que se mantienen estudiando en la institución o lograron concluir sus estudios.
  - Estudiantado desertor primer ingreso: incluye estudiantes que matricularon su primer semestre en la institución y no matricularon semestres posteriores, alcanzan al inicio del 2016 un mínimo de dos años consecutivos sin matricular.
  - Estudiantado desertor avanzado: incluye estudiantes que matricularon dos o más semestres, pero dejaron de hacerlo, al punto que a inicios del 2016 tenían un mínimo de dos años consecutivos sin matricular.
- Sociodemográficas: El índice de desarrollo social del distrito donde proviene el estudiantado (indicador socioeconómico), edad en años cumplidos de ingreso a la institución, sexo, nacionalidad y residencia igual o distinta en tiempo lectivo y no lectivo.
- Programa de estudio: El área de estudios donde ingresó, la sede de la universidad y la modalidad de estudio diurno o nocturno.
- Beneficio estudiantil: La obtención de beca al ingresar (exceptuando la beca más alta), obtención de la beca más alta al ingresar (Beca Mauricio Campos), obtención de financiamiento (tipo de beca que otorga un préstamo de dinero para financiar parcial o totalmente los costos de estudio), ingreso a primera opción de carrera, ingreso a carrera acreditada ya sea por el Sistema Nacional de Acreditación de la Educación Superior (SINAES) de Costa Rica, el Canadian Engineering Accreditation Board (CEAB) o la Agencia Centroamericana de Acreditación de Programas de Arquitectura y de Ingeniería (ACAAI).
- Historial académico: El tipo de colegio de donde proviene el estudiantado, la solicitud de algún tipo de adecuación para realizar la PAA y las notas en los componentes de la prueba de aptitud académica (PAA) del TEC: razonamiento matemático, razonamiento verbal y rendimiento del último ciclo de la educación secundaria. Hay que destacar que las notas de los componentes de la PAA fueron categorizados en tres grupos: el que obtuvo puntajes iguales o superiores a la media, el que obtuvo puntajes por debajo de la media y el que ingresó a la institución por un mecanismo distinto al examen de admisión (PAA). Se usaron las categorías en vez de los puntajes porque no todo el estudiantado realiza la prueba para ingresar.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

- **Historial universitario:** Incluye la nota promedio obtenida y la cantidad de créditos matriculados en el primer ciclo lectivo de la universidad. Estas dos variables se conocen cuando el primer ciclo lectivo de la universidad ha avanzado o bien cuando ha concluido. Los créditos matriculados se conocen al menos un mes después de que el primer ciclo lectivo ha iniciado, ya que el estudiantado puede matricular o retirar cursos durante ese periodo.

Los datos de estas variables se obtuvieron del SIGI de la Oficina de Planificación Institucional (OPI) del TEC. Este sistema permitió la consulta y extracción de los datos, los cuales están recolectados por distintas bases institucionales. La sistematización de esta información se realizó a partir de un procedimiento técnico, que fue validado por analistas de sistemas del TEC.

## Muestra

Esta investigación inició en el 2015 con la población del estudiantado que ingresó a un programa de grado de la universidad en el 2011, 2012 y 2013. La selección de estas tres cohortes consideró como primer criterio el parámetro establecido por la Comisión Nacional de Rectores de Costa Rica (CONARE), de contar con dos años consecutivos de no matrícula en la universidad para ser clasificado como estudiantado desertor. El segundo criterio fue que la población seleccionada tuviera características similares al futuro estudiantado de esta universidad, con la finalidad de usar estas variables en simulaciones que permitirán predecir las probabilidades de deserción en las cohortes siguientes. Un tercer criterio fue analizar esta problemática desde una perspectiva institucional, por lo tanto, se incluyó el estudiantado matriculado en las sedes de la Universidad (ubicadas en Cartago y San Carlos), en el Centro Académico de San José y en la sede interuniversitaria de Alajuela.

Un total de 5181 estudiantes cumplían con estas condiciones, sin embargo, se excluyeron algunos casos que carecían de información en las variables de estudio. También se eliminaron estudiantes que estaban en pasantías; de intercambio, dado que procedían de otras universidades y permanecían por un periodo inferior a un año, a los casos de revalidación y a quienes matricularon cursos bimestrales, los cuales no coincidían con la modalidad del programa de estudio que era semestral.

Como se muestra en la [Tabla 2](#), la muestra final estuvo compuesta por 4989 estudiantes, de los cuales 34,2% eran mujeres y 65,8% eran hombres, cuya edad promedio de ingreso al TEC fue de 19 años y el índice promedio del desarrollo social (IDS) del distrito de procedencia fue de 68,3% que corresponde a un nivel medio. En cuanto al colegio, el 61,1% procedían de colegios financiados con fondos públicos, el 15,5% de colegios subvencionados, el 22,5 de privados y un 0,9% de extranjeros.

doi: <http://dx.doi.org/10.15359/ree.23-1.5>URL: <http://www.una.ac.cr/educare>CORREO: [educare@una.cr](mailto:educare@una.cr)**Tabla 2:** Total de estudiantado matriculado en los programas de grado del Instituto Tecnológico de Costa Rica en el 2011, 2012 y 2013

Sedes	2011	2012	2013	TOTAL
Sede de Cartago	1 129	1 301	1 479	3 909
Sede interuniversitaria de Alajuela	9	53	79	141
Centro Académico San José	123	89	151	363
Sede de San Carlos	142	181	253	576
Total	1 403	1 624	1 962	4 989

**Nota:** Tomado del [TEC, Oficina de Planificación Institucional \(2018\)](#).

## Técnicas de análisis

En la fase explicativa se procesaron tres modelos de regresión multinomial para explicar la deserción joven y avanzada, siguiendo la metodología de [Cohen, Cohen, West y Aiken \(2003\)](#) con el paquete estadístico IBM SPSS, versión 19. En el primer modelo se ingresan las variables sociodemográficas y del programa de estudio para conocer su efecto directo sobre la deserción. Se ingresaron de primero, porque en modelos posteriores asumirán el rol de variables de control, ya que pueden ejercer influencia sobre restantes variables independientes y la dependiente de manera simultánea. En el segundo modelo se valora el efecto de las variables relacionadas con los beneficios estudiantiles obtenidos al ingresar a la universidad y las variables de historial académico sobre la deserción, independientemente de las variables sociodemográficas y del programa de estudio, como se mencionó anteriormente.

Por último, en el tercer modelo se analiza el efecto de los créditos matriculados y la nota obtenida por el estudiantado, luego de concluir el primer ciclo lectivo. Estas dos variables del historial universitario se agregan de último porque podrían estar influenciadas por las restantes variables independientes, que en este modelo ejercen el rol de control.

En la segunda fase predictiva se evaluó la capacidad para predecir la deserción de algoritmos generados con seis métodos: boosting trees, random forest, árboles de decisión, redes neuronales, support vector machines y regresión logística. Se evaluó cuál fue el mejor método predictor de la deserción en dos momentos distintos: cuando el estudiantado ingresa a la institución y cuando ha transcurrido el primer ciclo lectivo. Esto se debe a que las variables, nota de primer ciclo lectivo y créditos matriculados, se logran conocer solo después de que ha transcurrido el primer ciclo. La variable deserción se categoriza en dos: estudiantado desertor y no desertor, ya que en un principio ninguno de los métodos funcionó adecuadamente para discriminar entre estudiantado desertor de primer ingreso y desertor avanzado.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Para determinar la capacidad predictiva de las variables y elegir el mejor método de predicción se ejecutaron los siguientes pasos:

- Paso división de datos: Para efectos predictivos el archivo se debe dividir en al menos dos partes (Lantz, 2013), una para entrenamiento del modelo y la otra para evaluar su capacidad predictiva. La primera parte se conformó con un 75% de los sujetos elegidos de manera aleatoria y la segunda por el restante 25%.
- Paso entrenamiento: En esta fase se estimaron los parámetros de los seis métodos. Para este proceso se utilizó la librería caret (Classification and Regression Training) del software R. Con esta librería se calibran los modelos evaluando por defecto diferentes parámetros en cada método (Kuhn, 2008). Para determinar cuáles parámetros contribuyen a una mejor predicción se utiliza 5-fold cross-validation approach porque minimiza la varianza asociada al proceso de validación (Lantz, 2013). Este procedimiento consiste en dividir la muestra en 5 submuestras, y promediar los resultados de las predicciones. La elección de los parámetros idóneos de cada método se realiza con la estimación de rendimiento accuracy que se observa en la fórmula 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

Donde.

TP= verdaderos positivos

TN=Verdaderos negativos

FN= Faltos negativos

FP=Falsos positivos

- Paso validación: En este paso se evalúa la capacidad de predicción de los seis modelos estimados. Para ello inicialmente se obtienen las predicciones de los seis modelos con el 25% de la muestra que se había reservado. Posteriormente se estiman los indicadores de ajuste: probabilidad de detectar correctamente al estudiantado desertor, probabilidad de detectar correctamente al no desertor, sensibilidad y especificidad. La probabilidad de detectar correctamente al estudiantado no desertor y la especificidad se agregan simplemente para términos informativos, ya que se determina como mejor modelo aquel que maximiza simultáneamente la probabilidad de detectar correctamente al estudiantado desertor y la sensibilidad.

doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

## Resultados

En la [Tabla 3](#) se observan las estadísticas descriptivas de las variables en estudio. De acuerdo con los resultados obtenidos, la deserción estudiantil es de un 25.7%. Esta se compone de un 9.2 % de deserción de primer ingreso y 16.5% de avanza. Lo más relevante de la población de las 3 cohortes a nivel sociodemográfico y del programa de estudio es que la mayoría de estudiantes son hombres e ingresan principalmente a carreras del área ingenieril. En cuanto los beneficios estudiantiles, cabe destacar que tan solo un poco más de una quinta parte del estudiantado obtuvo la beca más alta llamada Mauricio Campos (Beca MC) y un 15% obtuvo otro tipo de beca. La otra modalidad de financiamiento de estudios fue concedida a un 10.4% de la población estudiantil. Una parte del estudiantado que obtuvo financiamiento también obtuvo otro tipo de beca. En el historial académico se incluyeron los tres componentes de la nota de admisión y se establecieron dos categorías, considerando como criterio el promedio: el grupo alto (notas superiores al promedio) y el grupo bajo (igual o superior al promedio). Finalmente, de las otras variables del historial académico sobresale que la mayoría proviene de colegios públicos académicos.

**Tabla 3:** Estadísticas descriptivas de las variables en la población de las cohortes 2011, 2012 y 2013

VARIABLES	Media	Desviación estándar
<b>Sociodemográficas</b>		
Índice desarrollo social	68.3	12.0
Edad	19	1.9
Estudiantado no desertor	74.3%	0.44
Estudiantado desertor primer ingreso	9.2%	0.29
Estudiantado desertor avanzado	16.5%	0.37
Sexo: Mujer	34.2%	0.47
Sexo: Hombre	65.8%	0.47
Nacionalidad: CR	97.8%	0.15
Nacionalidad: Otra	2.2%	0.15
Residencia igual	83.1%	0.37
Residencia distinta	16.9%	0.37

continúa...



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Variables	Media	Desviación estándar
<b>Programa de estudio</b>		
<b>Variable áreas</b>		
Área Básica	14.0%	0.35
Área Económica	19.2%	0.39
Área Educación	1.7%	0.13
Área Natural	13.7%	0.34
Área Ingeniería	51.4%	0.50
<b>Variable sede</b>		
Sede Alajuela	2.5%	0.16
Sede San Carlos	11.1%	0.31
Sede San José	6.7%	0.25
Sede Cartago	79.7%	0.40
<b>Variable jornadas</b>		
Jornada diurna	94.9%	0.22
Jornada nocturna	5.1%	0.22
<b>Beneficios estudiantiles</b>		
<b>Variable financiamiento</b>		
Estudiantado becado	15.0%	0.36
Estudiantado no becado	85.0%	0.36
Estudiantado financiado	10.4%	0.31
Estudiantado no financiado	89.6%	0.31
Estudiantado becado MC	21.5%	0.41
Estudiantado no becado MC	78.5%	0.41

continúa...



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Variables	Media	Desviación estándar
<b>Variable carrera</b>		
Carrera acreditada	66.6%	0.47
No carrera acreditada	33.4%	0.47
Ingreso primera opción	75.4%	0.43
Ingreso segunda opción	24.6%	0.43
<b>Historial académico</b>		
<b>Variable nota admisión</b>		
Examen mate alto	46.6%	0.50
Examen mate bajo	45.5%	0.50
Examen verbal alto	50.9%	0.50
Examen verbal bajo	41.2%	0.49
Nota colegio alta	48.1%	0.50
Nota colegio baja	44.0%	0.50
Ingreso sin examen	7.9%	0.27
<b>Variable colegio de procedencia</b>		
Extranjero	.9%	0.09
Privado	22.5%	0.42
Semipúblico académico	8.6%	0.28
Semipúblico no académico	6.9%	0.25
Público no académico	22.3%	0.42
Público académico	38.8%	0.49
<b>Variable adecuación en la PAA</b>		
Solicitud adecuación	2.1%	0.14
No solicitud adecuación	97.9%	0.14

**Nota:** Elaboración propia.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

En la **Tabla 4** se presentan los coeficientes de la regresión multinomial. En el modelo 1 se refleja el efecto de las variables sociodemográficas y del programa de estudio sobre la deserción. En el modelo 2 se analiza el efecto que ejerce sobre la deserción las variables relacionadas con los beneficios estudiantiles obtenidos al ingresar y el historial académico, independientemente de las variables sociodemográficas y del programa de ingreso, en tanto que, el modelo 3 se plantea para determinar la influencia de la nota promedio obtenida por los estudiantes durante el primer ciclo lectivo sobre la deserción.

**Tabla 4:** Coeficientes del modelo de regresión multinomial

Coeficiente	Modelo 1		Modelo 2		Modelo 3	
	Estudiantado desertor primer	Estudiantado desertor avanzado	Estudiantado desertor primer	Estudiantado desertor avanzado	Estudiantado desertor primer	Estudiantado desertor avanzado
<b>Sociodemográficas</b>						
Edad ingreso	1.20*	1.12*	1.190*	1.137*	1.108*	1.87*
Índice desarrollo social	1.01	1.01	0.996	1.008	1.001	1.10*
Sexo mujer	1.16	0.94	1.128	0.978	1.342	0.999
Nacionalidad CR	0.43*	0.67	0.556	0.627	0.352*	0.57*
Residencia igual	0.64*	0.61*	0.600*	0.581*	0.545*	0.55*
<b>Programa de estudio</b>						
Área Básica	1.23	1.36*	1.361	1.401*	1.048	1.39*
Área Económica	0.94	0.67*	0.868	0.612*	1.304	0.870
Área Educación	2.16*	1.81*	2.192*	1.525	4.113*	2.06*
Área Natural	1.02	0.92	0.968	0.905	1.386	1.170
Sede Alajuela	1.25	0.61	1.097	0.588	2.033	0.68
Sede San Carlos	0.96	0.81	1.212	0.838	1.725*	1.15
Sede San José	1.40	1.57*	1.268	1.495*	1.157	1.60*
Jornada diurna	0.79	1.01	0.722	0.901	1.193	1.463
<b>Beneficios estudiantiles</b>						
Becados			0.530*	0.669*	0.982	0.823
Financiados			0.148*	0.630*	0.472*	0.769
Beca MC			0.281*	1.075	0.602*	1.279*
Carrera acreditada			0.736*	0.793*	0.784	0.822*
Ingreso opción 1			0.913	0.845	1.317	0.986
Ingreso sin examen			0.652*	0.609*	1.037	0.701
Nota mate alto			0.741*	0.938	1.261	1.219*
Nota verbal alto			0.960	0.866	1.249	0.966
Nota colegio alta			0.914	0.836*	1.308	1.058

continúa...



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Coeficiente	Modelo 1		Modelo 2		Modelo 3	
	Estudiantado desertor primer	Estudiantado desertor avanzado	Estudiantado desertor primer	Estudiantado desertor avanzado	Estudiantado desertor primer	Estudiantado desertor avanzado
<b>Historial académico</b>						
Extranjero			1.556	0.674	7.227*	0.904
Privado			0.790	1.000	0.979	1.080
Semipúblico académico			0.474*	0.610*	0.495	0.668*
Semipúblico no académico			0.489*	0.836	1.519	1.258
Público no académico			0.717*	0.800*	1.065	0.933
Solicitud adecuación			1.285	1.057	0.659	0.816
<b>Historial universitario</b>						
Créditos					0.854*	0.902*
Nota					0.915*	0.963*
<b>Resultados de las pruebas estadísticas de los tres modelos</b>						
Chi square Test	186.89*		387.94*		1.885.384*	
Cox	0.037		0.075		0.315	
Nagelkerke	0.047		0.097		0.408	
McFadden	0.025		0.053		0.256	

**Nota:** Elaboración propia.

Los resultados del primer modelo sugieren que: el tener una residencia igual en tiempo lectivo y no lectivo disminuye la probabilidad de desertar (avanzado y de primer ingreso) mientras que ingresar con una edad más tardía a la universidad la incrementa. Ser de nacionalidad extranjera incrementa la probabilidad de desertar en los de primer ingreso. El área de educación muestra la mayor deserción de primer ingreso y avanzada, mientras que el área económica muestra la menor deserción. Por último, se observa que la sede de San José es la que presenta la mayor deserción avanzada.

Con el modelo 2 se concluye que el poseer beca, financiamiento, ingresar a una carrera acreditada, ingresar por un mecanismo distinto al examen admisión, o bien, el haber cursado estudios de secundaria en un colegio público no académico o semipúblico académico disminuye la probabilidad de deserción en estudiantado de primer ingreso y avanzado. Hay otras variables que influyen en solo uno de los dos tipos de deserción. Por ejemplo, quienes obtuvieron la beca Mauricio Campos al ingresar (es la mayor beca ofrecida por la institución) y quienes provienen de colegios semipúblicos no académicos tienden, en menor medida, a desertar jóvenes. Por otra parte, la nota del componente matemático del examen de admisión (PAA) y la nota de IV ciclo se correlacionan de manera negativa con la deserción avanzada. En el modelo 3 se demuestra que la nota obtenida durante el primer ciclo lectivo y la cantidad de créditos matriculados se relacionan con la deserción de primer ingreso y del estudiantado avanzado.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Estos resultados se obtienen a partir de los coeficientes del modelo que resultaron estadísticamente significativos. Cada coeficiente refleja el cambio en la razón entre la probabilidad de desertar y la probabilidad de no desertar, al cambiar en una unidad la variable explicativa, o al pasar de una categoría a otra en caso de que la variable sea cualitativa. Por ejemplo, en el modelo 1 la razón entre la probabilidad de ser estudiantado desertor avanzado y la probabilidad de no serlo cuando se estudia en el área de educación (codificada con 1) es 1.81 veces la razón entre esas probabilidades cuando se es estudiante del área de ingenierías (codificada con 0), o bien, se puede decir que la razón se incrementa en 81% cuando el estudiantado es parte del área de educación con respecto a quienes son de las ingenierías.

Además del modelo explicativo expuesto con anterioridad, se generaron seis modelos predictivos de la deserción al momento en que el estudiantado ingresa a la institución. Como variables predictivas se utilizaron las mismas del segundo modelo explicativo. En la [Tabla 5](#) se presentan los resultados de la capacidad predictiva de los seis modelos. El único que mostró cierta capacidad para detectar correctamente al estudiantado desertor es el “random forest” (lo detecta correctamente con una probabilidad de 0.88), sin embargo, apenas logra detectar un 2% de la deserción total, es decir, su sensibilidad es muy baja.

Luego de que el primer ciclo lectivo ha finalizado, se cuenta con dos variables que pueden ser trascendentales para predecir la deserción, como lo es el rendimiento obtenido en el primer ciclo lectivo y la cantidad de créditos cursados en el primer ciclo. Por esta razón se vuelven a evaluar los seis métodos incluyendo las dos variables mencionadas. Como se puede apreciar en la [Tabla 6](#), los modelos mejoran sustancialmente la capacidad de detectar correctamente al estudiantado desertor y la sensibilidad.

**Tabla 5:** Ajustes de los modelos antes de la primera matrícula

Capacidad de ajuste	Random Forest (mtry = 2)	Boosted Tress (n.tress = 50, interaction.depth = 3, shrinkage = 0.1 y n.minobsinnode = 10)	Tree (trials = 20, model = tree y winnow = TRUE)	Red Neuronal size = 1 y decay = 0.1	SVM sigma = 0,01401619 y C = 1	Logística decay = 0,1
Verdaderos positivos	0,88	0,67	0,63	0,00	0,70	0,51
Verdaderos negativos	0,75	0,75	0,75	0,75	0,75	0,75
Sensibilidad	0,02	0,04	0,07	0,00	0,02	0,06
Especificidad	1,00	0,99	0,09	1,00	1,00	0,98

**Nota:** Elaboración propia.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

**Tabla 6:** Ajustes de los modelos predictivos después del I ciclo lectivo

Capacidad de ajuste	Random Forest (mtry = 2)	Boosted Tress (n.tress = 50 interaction.depth = 3, shrinkage = 0.1 y n.minobsinnode = 10)	Tree (trials = 20, model = tree y winnow = TRUE)	Red Neuronal size = 1 y decay = 0.1	SVM sigma = 0,01401619 y C = 1	Logística decay = 0,1
Verdaderos positivos	0,83	0,73	0,69	0,70	0,83	0,73
Verdaderos negativos	0,81	0,82	0,82	0,82	0,80	0,82
Sensibilidad	0,34	0,41	0,39	0,41	0,31	0,39
Especificidad	0,98	0,95	0,94	0,94	0,98	0,95

**Nota:** Elaboración propia.

Al evaluar ambos indicadores de capacidad de ajuste, el “random forest” muestra la mejor combinación. Predice correctamente un 83% de lo que clasifica como estudiantado desertor (verdaderos positivos) y detecta un 34% (sensibilidad) de la totalidad. Las máquinas de soporte vectorial (svm) también predicen un 83% de estudiantado desertor (verdaderos positivos), sin embargo, muestran una sensibilidad un poco más baja. El “boosted trees” muestra una mayor sensibilidad puesto que detecta un 7% más de estudiantado desertor, pero con una probabilidad menor de detección correcta, la cual llega a 0.67. Los otros tres algoritmos presentan una sensibilidad y un porcentaje de verdaderos positivos más bajos que el random forest. Es importante resaltar que el random forest se obtiene a partir de una combinación de árboles, que de forma independiente clasifican a cada sujeto en desertor y no desertor. La clasificación final se genera a partir de una combinación de todas las clasificaciones obtenidas de cada árbol. Para este algoritmo la selección de los nodos de los árboles se obtiene de una selección aleatoria de dos variables.

## Conclusiones

El objetivo de este estudio fue desarrollar modelos explicativos y predictivos de la deserción estudiantil en los programas académicos de grado del Instituto Tecnológico de Costa Rica, con base en diferentes variables registradas en el Sistema de Indicadores de Gestión Institucional (SIGI). El modelo explicativo busca comprender la relación de la deserción con distintas variables recopiladas en un sistema de información de la universidad. Estas variables se clasificaron en cinco categorías: programa de ingreso, sociodemográficas, beneficios estudiantiles, historial académico e historial universitario después del primer ciclo lectivo en la universidad. El segundo modelo, predictivo, utiliza las variables explicativas del primer análisis para incorporarlas en algoritmos que permitan predecir al estudiantado desertor de manera anticipada.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Al finalizar el análisis explicativo se logró evidenciar relaciones significativas entre las variables de las cinco categorías mencionadas y la deserción. En el caso del programa de ingreso se observa que la deserción difiere de acuerdo con la sede y el área de estudio; el Centro Académico San José presenta mayor deserción avanzada que la de sede de Cartago, mientras que el área ingenieril muestra menor deserción de primer ingreso y avanzada que el área de educación, pero mayor deserción que el área económica.

El análisis de las variables sociodemográficas demostró que el tener una residencia igual en tiempo lectivo y no lectivo disminuye la probabilidad de desertar, lo cual coincide con otros estudios realizados en el TEC (Barquero, 1993; Chinchilla, 2013) y en otros contextos universitarios donde se reporta un mayor abandono en el estudiantado que procede del área rural (Castillo, 2010). Con la variable edad de ingreso a la universidad, se obtiene que por cada año de incremento en la edad ingreso, se aumenta la probabilidad de ser estudiantado desertor de primer ingreso o avanzado. Lo anterior coincide con los estudios de Brenes (2006), Castaño et al. (2006), Regueyra y Rojas (2013), y Yasmin (2013), quienes reportan resultados similares; en tanto que, Castillo (2010) reporta un estudio donde se disminuye la tasa de abandono al aumentar la edad en el contexto de una Open University. Por último, en la categoría sociodemográfica también se halló que el poseer nacionalidad costarricense incrementa la probabilidad de deserción de primer ingreso.

Al analizar las variables de la categoría beneficios estudiantiles se concluye que el brindar algún tipo de apoyo económico cuando se ingresa a la institución tiene un impacto positivo sobre la disminución de la deserción, esto ya ha sido señalado por otros estudios (Barquero, 1993; Castillo, 2010; Regueyra y Rojas, 2013). Una conclusión novedosa es que se encuentra más efectivo en términos de reducción de deserción brindar un préstamo o financiamiento que otorgar una beca. Esto puede estar sustentado en el hecho de que una beca no genera ningún tipo de repercusión económica, si se abandona la universidad, mientras que el préstamo genera un compromiso económico futuro que le puede mantener en la universidad. Al respecto, Donoso y Schiefelbein (2007, p. 16) señalan que desde el enfoque costo-beneficio cada estudiante analiza su capacidad para financiar sus estudios y toma una decisión de acuerdo con otras variables, tales como: créditos de largo plazo o subsidios parciales o totales, entre otras.

Otra de las variables de beneficio estudiantil relacionado, de manera inversa, con la deserción de primer ingreso y avanzado fue el estudiar una carrera acreditada. A nivel institucional, las carreras del TEC se acreditan a través SINAES de Costa Rica, el CEAB o la ACAAI; entidades que evalúan atributos y estándares de calidad relacionados con la formación profesional, que podrían contribuir a la permanencia de estudiantes por el prestigio de la carrera acreditada y el proceso de mejora en la que se ve inmersa la carrera mediante evaluaciones externas.

En cuanto a las variables de historial académico, se debe destacar que la prueba de admisión de la universidad evalúa las habilidades matemáticas y habilidades verbales básicas que requiere el estudiantado de primer ingreso. Uno de los principales hallazgos evidencia que las habilidades

doi: <http://dx.doi.org/10.15359/ree.23-1.5>URL: <http://www.una.ac.cr/educare>CORREO: [educare@una.cr](mailto:educare@una.cr)

matemáticas son un determinante de la deserción de primer ingreso, ya que quienes obtienen puntajes más altos en el componente matemático desertan en menor medida luego de concluir el primer semestre; por su parte, las habilidades verbales no muestran ninguna influencia sobre la deserción. Sin embargo, no todo el estudiantado que ingresa a la institución realiza la prueba de admisión; una parte ingresa por otros mecanismos (convenios o exención) y se ha logrado comprobar que quienes ingresan por estas dos opciones desertan en menor medida.

En el caso del proceso de exención, se trata de población estudiantil que se encontraba matriculada en otra institución de educación superior y por el proceso de exención se trasladaron al TEC, por lo que son estudiantes con alguna experiencia universitaria previa y probablemente mayor conocimiento de la carrera que desean estudiar. Por su parte, la población estudiantil que ingresa por convenios tiene la característica de haber concluido estudios a nivel técnico en colegios universitarios, lo que, a partir de algún convenio firmado entre ambas instituciones de enseñanza, les da oportunidad de ingresar a una carrera a nivel de grado en la misma área de estudio; por tanto, es una población estudiantil más segura de su carrera universitaria.

Finalmente, en la categoría de antecedentes académicos se encontraron evidencias de la relación entre el rendimiento obtenido por el estudiantado en el colegio (nota de educación diversificada) con la deserción. Específicamente, quienes obtienen puntajes más altos en sus notas de la educación diversificada, presentan menor deserción avanzada; lo que se sustenta en que “el desempeño académico en [la universidad está] fuertemente ligado a la formación básica con la que el estudiante ingresa a los estudios superiores”, como lo apuntan en sus investigaciones previas [Giuliodori, Gertel, Casini, y González \(2012, p. 17\)](#). De manera que, un mejor desempeño en su formación básica y en la universidad hace que las probabilidades de abandono se reduzcan.

Otra de las variables de historial académico que muestra relación con la deserción es el colegio de procedencia. Se observa que provenir de un colegio semipúblico, o bien de uno público no académico es un atributo que reduce la deserción. En Costa Rica los colegios semipúblicos cuentan con la ventaja de que el personal docente y administrativo es financiado por el Estado, pero laboran para una institución privada; en tanto que, en los colegios de modalidad técnica (no académicos), la formación es por tres años y se centra en desarrollar capacidades para que los grupos jóvenes se incorporen en el sector productivo con un nivel técnico. Estas características académicas de los colegios, en conjunto con la condición económica y atributos académicos del estudiantado, en definitiva, contribuyen a su permanencia en la universidad.

Por último, se analiza el rendimiento obtenido luego del concluir el primer ciclo lectivo sobre la deserción, independientemente de los créditos matriculados. Esta variable ha sido ya retomada por otras investigaciones ([Abarca y Sánchez, 2005](#); [Castaño et al., 2006](#); [Castillo, 2010](#); [Cruz, Gática, García y Hernández, 2010](#); [Merlino et al., 2011](#); [Regueyra y Rojas, 2013](#)). En el presente estudio sobresale el hecho de que el rendimiento obtenido durante el primer ciclo



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

lectivo se relaciona directamente con la deserción de primer ingreso y la deserción avanzada, conforme aumenta el promedio, disminuye la probabilidad de deserción. Este resultado debe contextualizarse a partir de la relación con las otras variables del modelo; además, de considerarse otros posibles factores como las experiencias emocionales positivas, que de acuerdo con otros estudios juega un importante rol en el desempeño académico y puede tener un considerable impacto en el conocimiento, comportamiento y el éxito final del estudiantado en el ámbito académico (Cruz et al., 2010).

Con respecto a los resultados obtenidos del enfoque predictivo se concluye que las variables recopiladas cuando el alumnado ingresa a la institución son insuficientes para predecir correctamente la deserción. Al finalizar el primer ciclo se pueden obtener dos variables adicionales que mejoran considerablemente la capacidad predictiva, como lo son: cantidad de créditos cursados y nota media de créditos cursados. Aun así queda mucho margen para mejorar la predicción, si se toma en cuenta que el algoritmo con mayor precisión para detectarla correctamente fue el "random forest" con una probabilidad de 0.83 de detectarla correctamente y una sensibilidad (porcentaje de estudiantado desertor detectado) de apenas 0.34. Este algoritmo se podría empezar a utilizar para detectar anticipadamente aquel estudiantado que dejará de matricular por un periodo de al menos dos años, considerando que una probabilidad de 0.83 de detectar correctamente es aceptable. El estudiantado que el algoritmo señala como posible desertor se podría intervenir anticipadamente para evitar que deje la institución.

Otros modelos predictivos de deserción universitaria, como los publicados por Delen (2010) y Pal (2012), utilizaron métodos similares de aprendizaje automático y lograron mejor sensibilidad y probabilidad de detección correcta del estudiantado desertor, por lo que cabe preguntarse a qué se deben esas diferencias. La primera posible respuesta son las variables que usaron. Al igual que en este estudio utilizaron variables de origen sociodemográfico, historial académico y habilidades académicas recopiladas previo al inicio de la universidad, rendimiento académico de los primeros cursos que llevaron en la universidad, programa de estudio, becas o beneficios económicos obtenidos para cursar el ciclo lectivo. No obstante, en términos de variables específicas hay algunas diferencias, por ejemplo, Delen (2010) cuenta con más variables de habilidades académicas, previas al inicio de la universidad como los diferentes rubros de medición del test SAT. Por su parte, Pal (2012) dispone de variables que miden la ocupación y escolaridad del padre y madre del estudiante, aunque a excepción de estas, las restantes que utiliza están incorporadas en el presente estudio.

Una segunda respuesta es la población de estudio, ya que la capacidad predictiva de la deserción de una variable puede depender de la cultura y condiciones de los países. En este caso ninguno de los estudios que mostraron mejor predicción se llevaron a cabo en América Latina o un contexto similar al presente. Otra variación con la población de estudio es el tipo de carrera donde se enfoca la investigación. Pal (2012) analiza únicamente estudiantes de ingeniería,

doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

mientras que el presente estudio maneja un espectro más amplio. Una tercera diferencia es la definición de deserción, para [Delen \(2010\)](#) y [Pal \(2012\)](#) el estudiantado desertor es aquel que no matricula luego del primer año de carrera, mientras que en este estudio se caracterizó como aquel con al menos dos años consecutivos sin matricular.

Una limitación en este estudio para obtener un modelo predictivo de mayor precisión y especificidad fue la carencia de otras variables en el sistema de información, que han mostrado alguna evidencia de relación con la deserción en otros estudios realizados en Costa Rica; por ejemplo, el interés del estudiantado en la carrera seleccionada. En esta investigación se incluyó una posible variable proxi, que indica si el estudiantado logró ingresar a su primera o segunda opción de carrera. Se parte del supuesto que quienes ingresan a la segunda opción pueden tener menor interés, no obstante, un estudiante puede tener amplio interés tanto en su primera como segunda opción; inclusive bajo interés de estudiar cualquiera de sus dos opciones; o bien, perder el interés posteriormente. Por esa razón habría que evaluar su interés en la carrera cuando matricula los cursos por primera vez y en próximas matrículas. Otro aspecto por indagar es la carrera que más deseaba estudiar. El hecho de que haya ingresado a su primera opción no significa que sea su mayor anhelo. Puede ser que haya elegido una carrera de menor demanda estudiantil para poder ingresar a la universidad y luego trasladarse a la carrera de su mayor interés. Ante estas posibilidades resulta necesario medir de manera más precisa el interés y vocación por la carrera seleccionada

El interés de estudiar en la universidad a la que ingresó es otra variable que requiere ser analizada con mayor precisión. De acuerdo con el estudio realizado por [Chinchilla \(2013\)](#) en el Tecnológico de Costa Rica, el 15,9% de estudiantes desertaron porque deseaban o preferían estudiar en otra universidad. La distancia para desplazarse al centro educativo y los horarios de las clases son dos de las posibles razones que incentivan la preferencia por otras instituciones. Por ello, se podría indagar cuando el estudiantado realiza su primera matrícula, si su principal opción de estudio era otra universidad y que tanto interés tiene de ingresar al TEC. Adicionalmente se deben medir los aspectos que motivan las preferencias por otras universidades. Dos de esos posibles aspectos son el tiempo de desplazamiento del hogar de residencia a la universidad y qué tanto le agradan los horarios de los cursos que se ofrecen.

Las condiciones socioeconómicas del hogar donde proviene el estudiantado y su condición laboral son otras de las variables que parecen tener amplio potencial predictivo de acuerdo con estudios previos realizadas en Costa Rica ([Chinchilla, 2013](#)). Además de lo mencionado, hay que tomar en cuenta aquellas variables que en estudios realizados en otros países han mostrado, de manera consistente, una fuerte correlación con la deserción, por ejemplo, el nivel educativo del grupo familiar del estudiantado, apoyo de la familia, actitudes y personalidad del sujeto, entre otras.



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Considerando estos antecedentes y hallazgos del estudio, la principal recomendación es recolectar información más específica del perfil de entrada del estudiantado, con la finalidad de mejorar el pronóstico de riesgo de deserción. Específicamente, de aquellas variables que han mostrado alguna relación con la deserción en Costa Rica, tales como: motivación para estudiar la carrera seleccionada, interés por estudiar en la universidad a la que ingresó, condición económica, la tenencia de hijos o hijas, la condición de recibir beca en el colegio, el apoyo financiero o moral de otras personas, entre otras. Sin embargo, antes de empezar a recopilar esas nuevas variables en el SIGI, es necesario discriminar las que tienen más potencial predictivo mediante una encuesta retrospectiva a estudiantado desertor y no desertor.

Finalmente, es recomendable que CONARE valore el criterio de reducir a dos semestres consecutivos (un año) sin matrícula en la universidad para considerar al estudiantado como desertor, lo que permitiría una intervención más oportuna para reducir el abandono. En este sentido, es conveniente un estudio sobre el tiempo de retorno a la institución de quienes han dejado de matricular, al menos, un semestre, para elaborar una propuesta más precisa y real del estudiantado desertor en las universidades públicas de Costa Rica.

## Referencias

- Abarca, A. y Sánchez, M. A. (2005). La deserción estudiantil en la educación superior: El caso de la Universidad de Costa Rica. *Actualidades investigativas en educación*, 5(especial), 1-22. doi: <https://doi.org/10.15517/aie.v5i4.9186>
- Barquero, J. A. (1993). *Deserción universitaria*. Cartago, Costa Rica: Instituto Tecnológico de Costa Rica. Recuperado de <https://repositoriotec.tec.ac.cr/handle/2238/690>
- Barquero, J. A. (1997). *Rendimiento académico, deserción y graduación en el ITCR*. Cartago, Costa Rica: Instituto Tecnológico de Costa Rica. Recuperado de <https://repositoriotec.tec.ac.cr/handle/2238/701>
- Brenes, M. I. (2005). *Deserción y repitencia en la educación superior universitaria de Costa Rica*. San José, Costa Rica: CONARE.
- Brenes, M. I. (2006). *Motivos e implicancias de la deserción en la cohorte de admitidos en 1996 a las instituciones de educación superior universitaria estatal*. San José, Costa Rica: CONARE.
- Castaño, E., Gallón, S., Gómez, K. y Vásquez, J. (2006). Análisis de los factores asociados a la deserción y graduación estudiantil universitaria. *Lecturas de Economía*, 65, 9-36. Recuperado de <http://aprendeenlinea.udea.edu.co/revistas/index.php/lecturasdeeconomia/article/view/2639/2099>



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

- Castillo, M. (2010). Deserción a nivel universitario. *Ensayos pedagógicos*, 5(1), 27-41. Recuperado de <http://www.revistas.una.ac.cr/index.php/ensayospedagogicos/article/view/4482/4311>
- Centro Interuniversitario de Desarrollo. (2006). *Repitencia y deserción universitaria en América Latina*. Santiago de Chile: Autor. Recuperado de <https://www.cinda.cl/download/libros/Repitencia%20y%20Deserci%C3%B3n%20Universitaria%20en%20Am%C3%A9rica%20Latina.pdf>
- Chinchilla, S. (2013). *Algunos datos sobre la deserción en el Instituto Tecnológico de Costa Rica en el 2012*. Cartago, Costa Rica: TEC. Recuperado de [https://repositoriotec.tec.ac.cr/bitstream/handle/2238/6782/estudio\\_%20deserci%c3%b3n\\_%202013.pdf?sequence=1&isAllowed=y](https://repositoriotec.tec.ac.cr/bitstream/handle/2238/6782/estudio_%20deserci%c3%b3n_%202013.pdf?sequence=1&isAllowed=y)
- Cohen, J., Cohen, P., West, S. G. y Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.
- Consejo Nacional de Rectores. (2008). *Memoria de la Cátedra UNESCO de Dirección Universitaria. Indicadores del sistema universitario público de Costa Rica en el ámbito académico*. San José, Costa Rica: Autor.
- Consejo Nacional de Rectores. (2016). *Plan nacional de la educación superior universitaria estatal 2016-2020. CONARE - OPES*. San José, Costa Rica. Recuperado de <https://www.conare.ac.cr/planes20/files/assets/basic-html/page1.html>
- Cruz, E. R., Gática, L., García, P. E. y Hernández, J. (2010). Academic performance, school desertion and emotional paradigm in university students. *Contemporary Issues in Education Research*, 3(7), 25-35. Recuperado de <https://files.eric.ed.gov/fulltext/EJ1072662.pdf>
- Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems*, 49(4), 498-506. doi: <https://doi.org/10.1016/j.dss.2010.06.003>
- 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*, 33(1), 7-27. doi: <https://doi.org/10.4067/S0718-07052007000100001>
- Giuliodori, R. F., Gertel, H. R., Casini, R. y González, M. V. (2012). *Estudio de la deserción en la Universidad Nacional de Córdoba (Argentina). Aplicación de modelos de duración*. Recuperado de <http://2012.economicsofeducation.com/user/pdfs Sesiones/063.pdf>
- Instituto Tecnológico de Costa Rica (2014). *Políticas específicas 2015*. Cartago, Costa Rica: Autor. Recuperado de <https://www.tec.ac.cr/reglamentos/politicas-especificas-2015>



doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

Instituto Tecnológico de Costa Rica, División de servicios estudiantes y académicos. Departamento de Orientación. (1980a). *Estudio deserción correspondiente al año 1979*. Cartago, Costa Rica: Instituto Tecnológico de Costa Rica. Recuperado de [https://repositoriotec.tec.ac.cr/bitstream/handle/2238/774/estudio de desercion correspondiente.pdf?sequence=2&isAllowed=y](https://repositoriotec.tec.ac.cr/bitstream/handle/2238/774/estudio_de_desercion_correspondiente.pdf?sequence=2&isAllowed=y)

Instituto Tecnológico de Costa Rica, División de servicios estudiantes y académicos. Departamento de Orientación. (1980b). *Estudio deserción correspondiente al I semestre 1978*. Cartago, Costa Rica: Autor. Recuperado de <https://repositoriotec.tec.ac.cr/handle/2238/773>

Instituto Tecnológico de Costa Rica, División de servicios estudiantes y académicos. Departamento de Orientación. (1981). *Estudio deserción correspondiente al curso lectivo 1980*. Cartago, Costa Rica: Instituto Tecnológico de Costa Rica. Recuperado de <https://repositoriotec.tec.ac.cr/handle/2238/2716>

Instituto Tecnológico de Costa Rica, Oficina de Planificación Institucional. (2014). *Compendio estadístico 2007-2013*. Cartago, Costa Rica: SIGI. Recuperado de [https://www.tec.ac.cr/sites/default/files/media/doc/compendio\\_estadistico\\_2007-2013\\_1.pdf](https://www.tec.ac.cr/sites/default/files/media/doc/compendio_estadistico_2007-2013_1.pdf)

Instituto Tecnológico de Costa Rica, Oficina de Planificación Institucional. (2018). *Bases de datos del Departamento de Admisión y Registro*. Cartago: Autor.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5). doi: <https://doi.org/10.18637/jss.v028.i05>

Lantz, B. (2013). *Machine learning with R*. Birmingham: Packet Publishing.

Merlino, A., Ayllón, S. y Escanés, G. (2011). Variables que influyen en la deserción de estudiantes universitarios de primer año. Construcción de índices de riesgo de abandono. *Revista Actualidades Investigativas en Educación*, 11(2), 1-30. doi: <https://doi.org/10.15517/aie.v11i2.10189>

Pal, S. (2012). Mining educational data to reduce dropout rates of engineering students. *International Journal of Information Engineering and Electronic Business*, 4(2), 1-7. doi: <https://doi.org/10.5815/ijieeb.2012.02.01>

Regueyra, G. y Rojas, C. (2013). Una mirada a la permanencia de la población estudiantil de trabajo social. *Revista Actualidades Investigativas en Educación*, 13(3), 1-25. Recuperado de <http://revistas.ucr.ac.cr/index.php/aie/article/view/12047/18275>

doi: <http://dx.doi.org/10.15359/ree.23-1.5>

URL: <http://www.una.ac.cr/educare>

CORREO: [educare@una.cr](mailto:educare@una.cr)

- Treviño, M., Ibarra, S., Castán, J., Laria, J. y Guzmán, J. (2013). A framework to avoid scholar desertion using artificial intelligence. En *Proceedings of the World Congress on Engineering*, 3, 1493-1497. Recuperado de [http://www.iaeng.org/publication/WCE2013/WCE2013\\_pp1493-1497.pdf](http://www.iaeng.org/publication/WCE2013/WCE2013_pp1493-1497.pdf)
- Universidad Nacional. (2008). *Estudios para enfrentar los retos del presente y los desafíos del futuro*. Heredia, Costa Rica: Autor. Manuscrito inédito
- Vélez, A. y López, D. F. (2004). Estrategias para vencer la deserción universitaria. *Educación y Educadores*, 7, 177-203. Recuperado de <http://www.alfaguia.org/alfaguia/files/1319735698estrategias%20para%20vencer%20la%20desercion%20universitaria%202004.pdf>
- Villamizar, G. A. y Pérez, L. (2011). Identificación de factores motivacionales y sociodemográficos de estudiantes desertores de la Facultad de Psicología de la Universidad Pontificia Bolivariana Bucaramanga. *Psicogente*, 14(25), 132-150. Recuperado de <http://publicaciones.unisimonbolivar.edu.co/rdigital/ojs/index.php/psicogente/article/download/414/411>
- Yasmin, D. (2013). Application of the classification tree model in predicting learner dropout behavior in open and distance learning. *Distance Education*, 34(2), 218-231. doi: <https://doi.org/10.1080/01587919.2013.793642>

