

UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E INDUSTRIAL/CENTRO DE POSGRADOS

MAESTRÍA EN MATEMÁTICA APLICADA

Tema: Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica

Trabajo de Titulación previo a la obtención del Grado Académico de Magister en
Matemática Aplicada

Modalidad de titulación Proyecto de Desarrollo

Autora: Ing. Cristina Nataly Vinueza López

Director: Ing. Edison Fernando Loza Aguirre PhD

Ambato-Ecuador

2021

APROBACIÓN DEL TRABAJO DE TITULACIÓN

A la Unidad de Titulación de la Facultad de Ingeniería en Sistemas, Electrónica E Industrial.

El Tribunal receptor de la Defensa del Trabajo de Titulación presidido por Ing. Elsa Pilar Urrutia Urrutia Mg., e integrado por los señores: Ing. Edith Elena Tubón Núñez Mg. e Ing. Washington Klever Medina Guerra Mg. designados por la Unidad de Titulación de la Universidad Técnica de Ambato, para receptor el Trabajo de Titulación con el tema: “Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica”, elaborado y presentado por la señorita Ing. Cristina Nataly Vinuesa López, para optar por el Grado Académico de Magister en Matemática Aplicada; una vez escuchada la defensa oral del Trabajo de Titulación el Tribunal aprueba y remite el trabajo para uso y custodia en las bibliotecas de la Universidad Técnica de Ambato.



Ing. Elsa Pilar Urrutia Urrutia Mg.
Presidente y Miembro del Tribunal de Defensa



Ing. Edith Elena Tubón Núñez Mg.
Miembro del Tribunal de Defensa



Ing. Washington Klever Medina Guerra Mg.
Miembro del Tribunal de Defensa

AUTORÍA DEL TRABAJO DE TITULACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el Trabajo de Titulación presentado con el tema: “Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica”, le corresponde exclusivamente a: Ing. Cristina Nataly Vinueza López, Autora bajo la Dirección de Ing. Edison Fernando Loza Aguirre PhD, Director del Trabajo de Investigación; y el patrimonio intelectual a la Universidad Técnica de Ambato.



Ing. Cristina Nataly Vinueza López

AUTORA



Ing. Edison Fernando Loza Aguirre PhD

DIRECTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que el Trabajo de Titulación, sirva como un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los Derechos de mi Trabajo de Titulación, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad Técnica de Ambato.



Ing. Cristina Nataly Vinueza López

C.C. 1804486650

ÍNDICE GENERAL

ÍNDICE DE TABLAS	vii
ÍNDICE DE FIGURAS	viii
AGRADECIMIENTOS	x
DEDICATORIA	xi
RESUMEN EJECUTIVO	xii
EXECUTIVE SUMMARY	xiv
CAPITULO I.....	1
EL PROBLEMA DE INVESTIGACIÓN.....	1
1.1. Introducción	1
1.2. Justificación	1
1.3. Objetivos	2
1.3.1. General	2
1.3.2. Específicos	2
CAPITULO II	3
ANTECEDENTES INVESTIGATIVOS.....	3
CAPITULO III.....	13
MARCO METODOLÓGICO	13
3.1. Ubicación	13
3.2. Equipos y materiales	13
3.3. Tipo de investigación	14
3.3.1. Investigación bibliográfica	14
3.3.2. Investigación de campo	14
3.3.3. Investigación descriptiva	14
3.4. Prueba de hipótesis.....	14
3.5. Población o muestra	14
3.6. Recolección de información	15
3.7. Procesamiento de la información y análisis estadístico	15
3.8. Variables respuesta o resultados alcanzados.....	17
CAPITULO IV.....	18

RESULTADOS Y DISCUSIÓN.....	18
4.1. Análisis descriptivo de las variables socioeconómicas de los estudiantes involucrados en la investigación.....	18
4.2. Análisis multivariante con regresión logarítmica para la obtención del modelo predictivo	22
4.3. Implementar el modelo predictivo para determinación de probabilidad de deserción de los estudiantes involucrados.....	29
4.4. Verificación de hipótesis	33
CAPITULO V	35
CONCLUSIONES, RECOMENDACIONES, BIBLIOGRAFÍA Y ANEXOS	35
5.1. Conclusiones	35
5.2. Recomendaciones.....	36
5.3. BIBLIOGRAFÍA.....	37
5.4. ANEXOS	40

ÍNDICE DE TABLAS

Tabla 1.- Variables dependientes del proyecto de investigación	16
Tabla 2.- Distribución porcentual en la muestra de las variables independientes.....	18
Tabla 3.- Coeficientes y valores de razón Odd del modelo de regresión logística 4	26
Tabla 4.- Comparación de cuatro modelos de regresión logística estudiados	28
Tabla 5.- Matriz de confusión de datos de entrenamiento	29
Tabla 6.- Matriz de confusión de datos de testeo	30
Tabla 7.- Valores de evaluación de la matriz de confusión de datos de testeo	30
Tabla 8.- Matriz de confusión de árbol de decisión	31
Tabla 9.- Valores de evaluación de la matriz de confusión de árbol de decisión	32

ÍNDICE DE FIGURAS

Figura 1. – Ubicación geográfica del Instituto Superior Tecnológico Luis A. Martínez Agronómico	13
Fuente: Google maps, 2020.....	13
Figura 2. – Metodología KDD	15
Fuente: Fayyad, Piatetsky-Shapiro y Smyth, 1996).....	15
Elaborado por: Vinueza, 2021 (Fuente: IST Luis A. Martínez Agronómico)	18
Figura 3. – Diagrama de caja y bigotes de la edad en función de la deserción estudiantil	20
Elaborado por: Vinueza, 2021.....	20
Figura 4. – Diagrama de caja y bigotes de la edad en función de la deserción estudiantil	20
Elaborado por: Vinueza, 2021.....	20
Figura 5. – Diagrama de barras de la deserción estudiantil del instituto	21
Elaborado por: Vinueza, 2021.....	21
Figura 6. – Resumen de modelo de regresión logística 1	22
Elaborado por: Vinueza, 2021.....	22
Figura 7. – Resumen de modelo de regresión logística 2	23
Elaborado por: Vinueza, 2021.....	23
Figura 8. – Resumen de modelo de regresión logística 3	24
Elaborado por: Vinueza, 2021.....	24
Figura 9. – Resumen de modelo de regresión logística 4	24
Elaborado por: Vinueza, 2021.....	24
Figura 10. – Análisis de significancia del modelo de regresión logística 4.....	25
Elaborado por: Vinueza, 2021.....	25
Figura 11. – Análisis ANOVA del modelo de regresión logística 4.....	26
Elaborado por: Vinueza, 2021.....	26

Figura 12. – Valores de razón Odd de las variables repitencia y carrera del modelo de regresión logística 4.....	27
Elaborado por: Vinueza, 2021.....	27
Figura 13. – Prueba de razón de similitud entre modelo 1 y modelo 4	29
Elaborado por: Vinueza, 2021.....	29
Figura 14. – Árbol de decisión podado a dos niveles	31
Elaborado por: Vinueza, 2021.....	31

AGRADECIMIENTOS

A Dios por todas las bendiciones que ha derramado sobre mí a pesar de todas mis equivocaciones

A mis padres porque su esfuerzo, amor y dedicación me han entregado todo lo que bueno que tengo

A mis hermanos, cuñada y sobrinos que me han llenado de cariño y apoyo

A las autoridades y personal del IST Luis A. Martínez Agronómico por su colaboración y ayuda en el desarrollo del proyecto

Al Ing. Edison Fernando Loza Aguirre PhD, director del proyecto quien siempre me supo guiar y además de su excelencia profesional en todo momento demostró su calidad humana

A todas aquellas personas que me prestaron su ayuda y consejos durante este largo camino

DEDICATORIA

A mis padres por todo el amor y apoyo que he recibido

A mi familia por todo el cariño y buenos consejos dados

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL/CENTRO DE POSGRADOS
MAESTRÍA EN MATEMÁTICA APLICADA

TEMA:

Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica

AUTORA: Ing. Cristina Nataly Vinueza López

DIRECTOR: Ing. Edison Fernando Loza Aguirre PhD

LÍNEA DE INVESTIGACIÓN:

- Diseño, materiales y producción

FECHA: 18 de enero de 2021

RESUMEN EJECUTIVO

En el presente proyecto se desarrolló un modelo de regresión logística para estimar la deserción estudiantil del Instituto Superior Tecnológico Luis A. Martínez Agronómico. Para lo cual se analizaron los datos de 849 estudiantes matriculados entre 2018 y 2020, estudiándose las variables independientes: género, estado civil, edad, carrera, repitencia, ocupación e ingresos económicos. Para desarrollar el modelo matemático se utilizó la metodología KDD, que permite generar información a partir de una base de datos con los registros a estudiarse. Dentro el período evaluado el 82,45 % de los estudiantes no desertaron y el 17,55% sí. Para el estudio se establecieron cuatro modelos de regresión logística, el primero incluye todas las variables independientes, siendo significativa solamente la variable 'carrera. Se eliminaron las variables 'edad' y 'género' (mayor valor p) y se obtuvo el modelo de regresión logística 2, las variables 'repitencia' y 'carrera' se consideran significativas.

Posteriormente, se eliminaron las variables con mayor valor p , 'estado civil' e 'ingresos económicos', obteniéndose el modelo de regresión logística 3, las variables 'repitencia' y carrera continúan siendo las únicas significativas. Por último, se escoge el modelo de regresión logística 4, el cual solamente incluye las variables 'carrera' y 'repitencia', que fueron las únicas significativas. Se rechazó la hipótesis nula, porque el coeficiente β_1 y β_2 de las variables carrera y repitencia son diferentes de cero. El modelo de regresión logística 4 clasificó correctamente el 83 % de los datos de entrenamiento y el 79 % de los datos de testeo. Adicionalmente, se determinó un modelo de predicción con árboles de decisión, que estableció como variable explicativa 'carrera'. El valor $F1_Score$ del modelo de regresión logística 4 fue mayor que el valor del $F1_Score$ del modelo con árbol de decisión.

Descriptor: deserción estudiantil, modelo matemático, regresión logística, predicción con árbol de decisión, educación superior, metodología KDD, datos entrenamiento, datos testeo, matriz de confusión, valor $F1_Score$

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL/CENTRO DE POSGRADOS
MAESTRÍA EN MATEMÁTICA APLICADA

THEME:

Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica

AUTHOR: Ing. Cristina Nataly Vinueza López

DIRECTED BY: Ing. Edison Fernando Loza Aguirre PhD

LINE OF RESEARCH:

- Design, materials and production

DATE: January 18th, 2021

EXECUTIVE SUMMARY

In this research, a logistic regression model was used to estimate student dropout from the IST Luis A. Martínez Agronómico. The data of 849 students registered in the institute between 2018 and 2020 was used to build the model. The independent variables considered for the model were: gender, marital status, age, career, repetition, occupation and economic status. We used the KDD methodology to estimate the mathematical model, which allows generating information from a database with the records to be studied. In the evaluated period, 82.45% of the students did not dropout but 17.55% did it. In the study, four logistic regression models were established, the first one includes all the independent variables but only the 'career' variable was significant. The 'age' and 'gender' variables were eliminated (higher p-value) for generating a second logistic regression model, where the 'repetition' and 'career' variables were considered significant. Subsequently, the highest p-value variables, 'marital status' and 'economic status' were eliminated for obtaining a third logistic

regression model wherein the ‘repetition’ and ‘career’ variables were the only significant ones. Finally, it was chosen the logistic regression model 4, which only includes the career and repetition variables as the only significant ones. The null hypothesis was rejected because the coefficients β_1 and β_2 of the variables ‘career’ and ‘repetition’ aren’t zero. The logistic regression model 4 correctly classified 83% of the training data and 79% of the test data. Additionally, we build a prediction model based on decision trees, which established ‘career’ as a unique explanatory variable. The F1_Score value of the logistic regression model 4 was higher than the F1_Score value of the decision tree model.

Keywords: student dropout, mathematical model, logistic regression, decision tree prediction, higher education, KDD methodology, train data, test data, confusion matrix, F1_Score value

CAPÍTULO I

EL PROBLEMA DE INVESTIGACIÓN

1.1. Introducción

La presente investigación pretende dotar de una herramienta que permita predecir el riesgo de deserción de estudiantes que cursan una carrera de nivel tecnológico superior en una institución del centro del país. Con este fin, se utilizarán modelos predictivos y técnicas de minería de datos para determinar patrones de comportamiento de los estudiantes, que determinen su condición de potencial desertor, asociándole un índice de deserción como probabilidad de abandono del sistema educativo. Dicha información podrá ser utilizada por las instituciones de educación superior del país para tomar medidas que eviten o reduzcan el nivel de deserción al mínimo posible. El modelo predictivo se desarrollará mediante un Modelo Multivariante con Regresión Logarítmica, ya que los modelos de elección discreta son bastantes apropiados para analizar los factores determinantes de la probabilidad de un suceso como el que se pretende estudiar. La presente investigación es de tipo caso de estudio, ya que estudia un sujeto o una realidad de carácter específico. Cabe recalcar que los estudios de casos se utilizan especialmente cuando las preguntas "cómo" o "por qué" se plantean, el investigador tiene poco control sobre los eventos, y cuando se investiga un fenómeno contemporáneo dentro de su contexto de la vida real, considerando que los límites entre el fenómeno y el contexto no son claramente evidentes (Yin, 2003).

1.2. Justificación

La presente investigación se realiza para diseñar un modelo matemático predictivo, que permita identificar de forma temprana los casos de estudiantes que presenten mayor probabilidad de desertar, para que la institución pueda aplicar medidas preventivas que evitan dichos casos de deserción. El diseño de obtuvo mediante un análisis multivariable con regresión logarítmica. La disminución de la tasa de deserción es de vital importancia en las instituciones de educación superior, ya que representa un indicador de la calidad educativa evaluado por las instancias superiores.

En la educación, la deserción se refiere al abandono prematuro de un programa de estudios antes de alcanzar un título o grado, considerando un tiempo suficientemente

largo como para descartar la posibilidad de que el estudiante retome sus estudios (Himmel,2002). La deserción estudiantil constituye un grave problema en la educación de nivel superior, técnica, tecnológica o universitaria, ya que tiene una incidencia negativa sobre los procesos políticos, económicos, sociales y culturales del desarrollo nacional (Sopalo, Guevara y Burbano, 2020). Además, constituye un grave problema para la institución de educación superior, ya que el nivel deserción constituye el indicador “Tasa de retención” del Modelo de evaluación institucional para los institutos superiores técnicos y tecnológicos del proceso de acreditación establecido por el Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES, 2020). De allí la importancia de establecer una alternativa que les permita identificar los casos de estudiantes próximos a retirarse e implementar acciones preventivas para evitarlo.

1.3. Objetivos

1.3.1. General

- Diseñar un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica.

1.3.2. Específicos

- Realizar un análisis descriptivo de las variables socioeconómicas de los estudiantes de la institución de educación superior tecnológica involucrados en la investigación.
- Aplicar el modelo de análisis multivariante con regresión logarítmica para la obtención del modelo predictivo.
- Implementar el modelo predictivo para determinación de la probabilidad de deserción de los estudiantes involucrados.

CAPITULO II

ANTECEDENTES INVESTIGATIVOS

1. En el trabajo de Vicente y Cuji (2020) con el tema “Aplicación de la técnica de minería de datos para la predicción de la deserción estudiantil universitaria” realizado en la Universidad Técnica de Ambato, se propone el uso de la minería de datos para predecir y evitar en lo posible la deserción estudiantil. Utilizando la metodología Knowledge Discovery in Databases (KDD), con el uso del programa R y la función Loggit (función matemática de regresión logística) se creó un modelo de regresión logística y se realizaron las pruebas necesarias para verificar el correcto funcionamiento del modelo, encontrando a las variables relacionadas con las calificaciones de los niveles segundo, cuarto y quinto como las más influyentes del modelo. El modelo logró una eficiencia del 98,59%, con lo que el modelo se considera válido.

2. Según Frutos y Buenaño (2017) en su trabajo de investigación “Análisis de la tasa de retención y su incidencia en la detección de patrones de deserción estudiantil en la Universidad Técnica de Ambato”, la deserción estudiantil es un problema social que no es ajeno a ninguna institución de educación superior a nivel nacional. En su investigación, el autor analiza el problema usando técnicas y algoritmos de Minería de Datos para identificar patrones de interés, aplicando la metodología KDD que sigue un proceso de Selección, Preprocesamiento, Transformación, Minería de Datos e Interpretación de resultados. Se concluyó que el análisis de la tasa de retención es muy importante en la institución y la detección de patrones de deserción estudiantil permitirá a las autoridades de la institución la toma de decisiones oportunas para prevenirla.

3. Sopalo, Guevara y Burbano (2020) en su investigación “Análisis de los factores que inciden en la deserción estudiantil de los niños, niñas y adolescentes ecuatorianos en el periodo 2009-2019” realizada en la Escuela Politécnica Nacional, indica que la deserción estudiantil constituye un problema que afecta

negativamente al progreso económico, político, social y cultural de una sociedad. La investigación tuvo como objetivo estudiar cómo afectan las características individuales y contextuales de los niños, niñas y adolescentes ecuatorianos a la deserción estudiantil. Para este fin, se utilizó un modelo de respuesta binaria Probit y como fuente de datos se consultó la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU) y el Archivo Maestro de Instituciones Educativas (AMIE) en el periodo del 2009 - 2019. En conclusión, se verificó que ciertas características individuales como 'ser mujer' y 'vivir en el área rural' aumentan la probabilidad de deserción. Por otro lado, se encontró que, un mayor ingreso familiar asegura la permanencia del individuo en el sistema educativo. Por su lado, las características contextuales mostraron que la probabilidad de deserción aumenta si los estudiantes se encuentran en una 'jurisdicción bilingüe'. Por último, se observó que un aumento de estudiantes por centro educativo por cantón aumenta la posibilidad de abandono estudiantil.

4. Segura y Loza (2017) tuvieron como principal objetivo de investigación en su trabajo "Uso de árboles de decisión para predecir el rendimiento académico en función de factores socioeconómicos" determinar si los factores socioeconómicos afectan la educación y logros de los estudiantes de una Escuela Politécnica. Para ello, consideraron los datos socioeconómicos y académicos correspondientes a más de diez años de registros. Posteriormente utilizaron algoritmos de clasificación y técnicas de aprendizaje automático para determinar qué factores son los más influyente en el rendimiento académico. Se realizó un estudio comparando diferentes métodos de árboles de decisiones. Al final los autores reportaron que las variables 'beca académica', 'edad', 'provincia' y 'título de escuela secundaria' influyen en el rendimiento académico de los estudiantes. Los resultados de la investigación constituyen información importante para directores académicos y trabajadores sociales involucrados en la tarea de mejorar las condiciones de estudiantes y proporcionándoles a todos los medios para el éxito.
5. González y Arismendi (2018) en su trabajo "Deserción Estudiantil en la Educación Superior Técnico Profesional: Explorando los factores que inciden en alumnos de primer año", buscaron identificar las variables que expliquen la deserción de los estudiantes usando los datos de la cohorte 2014-2016. La muestra

correspondió a 1876 estudiantes de primer año de un instituto técnico-profesional de la región de Valparaíso y los datos fueron procesados con el software R, aplicando un modelo lineal generalizado con distribución de errores binomial y función de vínculo Logit. Los resultados mostraron que los factores con mayor significancia estadística corresponden al ‘género’, ‘año de egreso de enseñanza media’ y ‘jornada de estudio’. Por último, es importante precisar que esto dependerá de la escuela de formación, pues el modelo varía de acuerdo con las particularidades de cada campo de formación.

6. El trabajo de Pérez, Escobar, Toledo, Gutiérrez y Reyes (2018), con título “Modelo de predicción de la deserción estudiantil de primer año en la Universidad Bernardo O’ Higgins” tuvo por objetivo diseñar un modelo predictivo de retención de los estudiantes de primer año de la Universidad Bernardo O’ Higgins - UBO (Santiago de Chile), identificando que las variables de entrada, ya sean académicas, sociales o familiares, se revelan significativas para el análisis. A partir de la sistematización de antecedentes socioeducativos de los estudiantes de las cohortes 2014 y 2015 disponibles en los sistemas informáticos de la universidad, se consolidó una matriz tripartita con los datos asociados a las variables que emergieron del análisis de los referentes consultados. Posteriormente el análisis bivariado permitió determinar diecisiete variables, significativamente asociadas con la deserción estudiantil; logrando a su vez identificar las relaciones de dependencia con el abandono de estudios. El modelo multivariado predijo en un 86,4 % la conducta de abandono señalando siete variables categóricas independientes que, al final se develan como factores relevantes del modelo predicción.
7. Albarrán (2019) en su investigación “La deserción estudiantil en la Universidad de Los Andes (Venezuela)” presenta un análisis para describir la deserción estudiantil en la Universidad de Los Andes (ULA-Venezuela). Para ello consultó varias fuentes bibliográficas y, debido a no contar con datos oficiales que midieran el fenómeno, usó información extraída de estimaciones de organizaciones no gubernamentales o diarios en línea. Los hallazgos revelan que 65% de los

estudiantes podrían dejar la institución por la falta de oportunidades laborales futuras, aumento de sus gastos personales y académicos, interrupción de las labores académicas por continuas protestas sociales, carencia de recursos económicos, baja formación escolar secundaria y desmotivación, principalmente. El estudio demostró que la deserción en la ULA es un fenómeno que se incrementó a medida que la crisis económica del país se agravó. Por lo tanto, se propuso que se debe brindar apoyo motivacional y académico a los estudiantes para que puedan permanecer en la institución.

8. Cuji, Gavilanes y Sánchez (2017) en su investigación “Modelo predictivo de deserción estudiantil basado en árboles de decisión” presentan la construcción de un modelo predictivo de deserción estudiantil con el objetivo de pronosticar la probabilidad de que un estudiante abandone su programa académico, a través de técnicas de clasificación, basadas en árboles de decisión. Se utilizó la metodología KDD en sus cinco etapas: selección, procesamiento, transformación, minería de datos y evaluación. Aplicando el algoritmo, Classification and Regression Tree (CART) de la herramienta R, se construyó un árbol con cuatro niveles de profundidad y mismo número reglas, que permite evaluar a los posibles desertores. Por último, los autores concluyeron que las variables ‘nivel’ y ‘notas’ tienen mayor influencia en la deserción.

9. En el trabajo de Poveda, Poveda y España (2020) se identificaron los factores que influyen sobre la deserción universitaria de los estudiantes de la Universidad Mayor, Real y Pontificia de San Francisco Xavier de Chuquisaca. El estudio se enfoca en entender mejor la deserción estudiantil a partir del análisis de ecuaciones estructurales, apoyado en el software SmartPLS versión 3. La población de referencia para el estudio fue el total de estudiantes matriculados al inicio de la gestión 2017 (46.347 estudiantes). Se realizó un muestreo probabilístico aleatorio para población finita conocida al 97% de confianza y 3% de error, con un total de 2.216 estudiantes encuestados (femenino 1.079, masculino 1.137). El análisis a través de ecuaciones estructurales determinó que los factores ‘Económico’ y ‘Familiar’ resultan ser los que tienen la mayor

influencia en el abandono estudiantil. Complementariamente se identificaron otros factores, como el de 'Motivación', 'Salud', 'Social' y 'Académico'. Sobre la base de los resultados, se recomienda realizar investigaciones complementarias de cada factor identificado.

10. Tapasco, Ruiz, Osorio y Ramírez, (2018) en su trabajo "Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión" estudiaron la incidencia que distintas variables institucionales del proceso de admisión tienen sobre el riesgo de deserción estudiantil. La población del estudio fueron los estudiantes matriculados en la Universidad de Caldas para las cohortes 2010-2012. Se utilizaron modelos logísticos multivariados y técnicas de optimización. Los modelos obtenidos arrojan que algunas variables del proceso de admisión, susceptibles de ser intervenidas institucionalmente, como el 'puntaje de admisión', 'ingreso por segunda opción', 'programas de cupos especiales' y 'costo de matrícula' tienen una incidencia estadísticamente significativa en el riesgo de deserción, influencia que varía entre facultades. En conclusión, para atenuar la deserción estudiantil e incrementar el poder predictivo del desempeño futuro del estudiante, se pueden contemplar acciones como el repensar el ingreso por segunda opción, el sostener las políticas de ingreso por méritos académicos y de acción afirmativa y el modificar las ponderaciones actualmente asignadas a las áreas evaluadas en el proceso de admisión.

11. Campoverde (2019) en su proyecto titulado "Deserción Estudiantil en el Instituto Tecnológico Superior de Fútbol en Guayaquil, 2019" se enfocó en la teoría de las estrategias humanas, que estudia al hombre real en su existencia genuina y desnuda, para estudiar el alto nivel de deserción estudiantil que existe en los estudiantes del Instituto Tecnológico Superior de Fútbol de Guayaquil. La población estudiada estuvo compuesta por 300 estudiantes y la muestra fueron 60 estudiantes, siendo un muestreo no probabilístico intencional. Se aplicó como instrumento un cuestionario con 20 preguntas, con un nivel de confiabilidad de 0.835 y cuya validez de contenido se evidenció a través del juicio de aprobación de tres expertos. Se concluyó que el 61.70% de alumnos se ubican en el nivel

regular de la dimensión Académica o Escolar correspondiente a la variable ‘deserción estudiantil’, rechazando la hipótesis de investigación. El estudio concluye que la deserción estudiantil es regular.

12. Pérez (2020) presenta los resultados de un estudio de comparación de técnicas para apoyar la identificación de deserción estudiantil a partir del registro académico de los estudiantes de la facultad de Sistemas de una Universidad en Colombia. El registro académico se estableció para un periodo de 7 años. Se utilizaron métodos como Árboles de decisión, regresión logística y Naive Bayes, para lograr establecer la mejor técnica de detección de desertores. Adicionalmente, la herramienta Watson Analytics de IBM fue utilizada para comparar su usabilidad y precisión para un usuario no experto. Los resultados experimentales obtenidos mostraron que las mejores áreas bajo la curva (AUC) se lograron por medio de Bosques Aleatorios, desde el tercer semestre de inscrito se obtuvo 0,91 AUC hasta el último semestre, donde el modelo dio 0,97 AUC. Cuatro características fueron utilizadas (‘Semestres’, ‘Prom SE’, ‘Falló SE’, ‘GPA’) para lograr esta precisión.
13. Según Camacho, Montalvo y Galezo (2019) en su investigación “Determinantes de la deserción estudiantil en estudiantes universitarios”, la deserción estudiantil es un problema significativo para el sistema educativo. Su trabajo se enfocó en un estudio descriptivo, correlacional que caracteriza a los 16.927 estudiantes de pregrado matriculados en la Universidad de Cartagena en diversas cohortes. Se establecieron los posibles determinantes en el riesgo de deserción y patrones temporales de deserción. Los datos se consultaron en la plataforma de SPADIES. La muestra de 1547 estudiantes desertores de la Universidad de Cartagena, fue proporcional a cada programa y se utilizó un instrumento de 22 preguntas, elaborado por los investigadores. Se encontró que en los últimos años la tasa de deserción en la institución había disminuido y estaba en correspondencia con el comportamiento de la tasa promedio de deserción a nivel nacional. Según el estudio, los hombres y los estudiantes de los programas de la modalidad a distancia presentaban más riesgo de desertar. Además, quienes ingresan con una clasificación baja en las pruebas Saber 11 presentaron un mayor riesgo de deserción en los primeros períodos de estudio.

14. Jurado y Plata (2019) en su proyecto de titulación “Diseño de un modelo predictivo de la deserción estudiantil de postgrado en una institución de educación superior” indica que la deserción estudiantil es un problema significativo en la unidad de postgrado objeto de estudio del proyecto, problema que afecta a varias áreas institucionales como: financiera, estándares de calidad y al ámbito académico. Por ello se requería diseñar un modelo predictivo de la deserción estudiantil, considerando cuatro modelos de clasificación que son: Potenciación del Gradiente Estocástico (GBM), Bosques Aleatorios (RF), Redes Neuronales (NNET) y Regresión Logística (GLM); además, se utilizó la metodología KDD cuyas etapas permiten la implementación de los modelos estadísticos en conjunto con técnicas computacionales para obtener un óptimo desempeño. Para definir el mejor modelo, se evaluó cuantitativamente cada uno de ellos por medio de la matriz de confusión, sus indicadores de desempeño, la curva ROC y el AUC. Se concluye que los modelos Bosques Aleatorios y Redes Neuronales se destacaron más, pero se escogió el modelo Bosques Aleatorios por su facilidad de uso.

15. Según Rodríguez (2016) en su investigación “Determinantes de la deserción escolar en la educación secundaria asociados a las características de las instituciones en el Ecuador”, en el país la educación se ha convertido en un eje fundamental de la política pública por lo que se ha invertido fuertemente en equipamiento, readecuación y construcción de centros educativos. No obstante, la limitada capacidad del sistema educativo para asegurar la permanencia de los alumnos en el sistema sigue siendo un problema latente en el Ecuador. En dicho sentido se orientó la investigación en determinar los factores relacionados con las características de las instituciones educativas que se asocian con la reducción o impulso del abandono escolar de la secundaria. Se utilizó un modelo de mínimos cuadrados ordinarios (MCO) considerando una estructura de clusters a nivel de circuito educativo, utilizando el AMIE. Los resultados concluyeron que las características de las instituciones educativas relacionadas con la infraestructura, equipamiento y características de los docentes; están relacionadas con la deserción escolar en la secundaria y presentaron diferencias entre ‘área’, ‘género’ y ‘subnivel educativo’.

16. En el trabajo de Reyes, Rosero y Medina (2019) con tema “Deserción escolar en estudiantes de 5 a 17 años en las Provincias de Pichincha y Guayas en el período 2013-2017” se identificaron los factores que inciden en la deserción escolar en estudiantes señalados, en las provincias de Pichincha y Guayas, utilizando la base de datos de la ENEMDU. Se determinó la probabilidad de deserción escolar mediante la realización de un pseudo panel, utilizando el modelo de probabilidad de efectos fijos, para identificar en que área y en qué nivel de educación existe mayor deserción. Se encontró como resultado relevante que la población estudiantil tiene una tendencia a desertar en los años más altos de educación como en el primer año de Bachillerato General Unificado (BGU). Además, el estudio de la cantidad de instituciones y el sostenimiento al que pertenecen tanto en Pichincha y Guayas determinó que mientras más oferta educativa exista mayor asistencia a clases existirá. Otros resultados indican que las mujeres tienen mayor probabilidad de deserción debido a factores como: falta de recursos económicos y búsqueda de trabajo a temprana edad. Adicionalmente, se encontró que la provincia con mayor deserción es Guayas. Finalmente, se identificó que en el área rural es donde más desertan los jóvenes, ya que no tienen los recursos económicos suficientes para completar sus estudios.

17. Machado y Castillo (2019) en su proyecto de titulación “Modelado del tiempo de estancia estudiantil en la FICFM de la Universidad Central del Ecuador” establecieron un modelo de probabilidad no paramétrico bi-dimensional para cuantificar la probabilidad de deserción de los estudiantes de la Facultad de Ingeniería, Ciencias Físicas y Matemática (FICFM) de la Universidad Central del Ecuador con base en las variables aleatorias tiempo de estancia y porcentaje de créditos aprobados, utilizando la información de registros de los últimos años que se manejan al interno de la Facultad. Adicionalmente, se determinaron modelos no paramétricos para estimar la probabilidad de salida de un alumno perteneciente a una población universitaria en base de modelos de supervivencia y demográficos; y también, aplicaron el modelo para estimar la dinámica de decremento de la población de estudiantes de la FICFM de la Universidad Central del Ecuador. Las variables aleatorias estudiadas fueron ‘tiempo de estancia’ y

‘porcentaje de créditos aprobados’; en base a la información de vectores de datos, y por medio del software libre R, se obtuvo la función de densidad conjunta de las variables ‘tiempo de estancias’ y ‘porcentaje de créditos aprobados’, utilizando el método de interpolación de densidad por Kernel. Con esta información se calculó el valor de la probabilidad de deserción-salida estudiantil por medio del uso de métodos numéricos que permitan calcular el volumen bajo la superficie dada por la densidad conjunta y el sector o área donde se ha definido a dicha densidad; considerando el hecho de que existen datos censurados y truncados. Con los resultados obtenidos se aplicaron los modelos supervivencia y demográficos para analizar la dinámica de decremento de la población de estudiantes de la FICFM de la UCE por medio de tablas de permanencia.

18. Según Castro (2020) en su investigación “Minería de datos educacionales para determinar de forma temprana la deserción estudiantil”, la deserción universitaria es uno de los principales problemas que afrontan las instituciones de educación superior y los sistemas de educación. Por ello se desarrolló una aplicación web, para determinar factores de riesgo de estudiantes propensos a desertar de manera permanente de una carrera universitaria. El aplicativo permitió la identificación de estudiantes en peligro de abandono de manera temprana y sus indicadores de deserción a través un dataset educativo. Además, mediante el planteamiento de un caso práctico se realizaron experimentos usando minería de datos y algoritmos de clasificación bayesianos y árboles de decisión usando la aplicación WEKA. Se encontró que los algoritmos de clasificación J48, RandomTree y BayesNet fueron los óptimos para inferir los factores más relevantes para determinar la deserción estudiantil. Al igual que en otras investigaciones, se comprobó que la deserción se produce en los primeros cuatro niveles y con una probabilidad de 88.3%.

19. En el trabajo de Vizúete y Chasiluisa (2017) con el tema “Deserción escolar de los estudiantes del colegio técnico Vicente Rocafuerte de la ciudad de Quito de la especialidad de mecánica automotriz y la repercusión en el ámbito laboral (año 2017-2017)” se investigó la deserción escolar en los estudiantes del colegio Técnico Vicente Rocafuerte en el periodo escolar 2017-2017, con el fin identificar los factores que influyen en la deserción escolar y su influencia el ámbito laboral.

Se utilizó una investigación científica basada en bibliografía enfocada en la socio-pedagogía, además se aplicó un cuestionario a los estudiantes de la institución. Se concluyó que la deserción escolar repercute e influye en la mayoría de los casos de manera negativa, y que este fenómeno afecta especialmente a los chicos que están en segundo y tercer año de bachillerato.

20. Zambrano, Rodríguez y Guevara (2018) en su revisión bibliográfica “Deserción estudiantil en las universidades del Ecuador y América” analizaron las causas y factores de deserción universitario y su influencia en los niveles de deserción en las universidades de Ecuador y América Latina, para identificar los factores que interaccionan e inciden en el aumento de la deserción. En el estudio se encontró que aspectos como el socioeconómico, educativo y psicológico influyen significativamente en esta problemática siendo los que representan mayor relevancia en los artículos, tesis y libros citados bibliográficamente en dicha revisión. Adicionalmente se logró identificar que a través de estrategias de retención y programas de seguimiento en los estudiantes durante el pregrado han disminuido significativamente la deserción universitaria en otros países. Los programas permiten tener indicadores donde los elementos de mayor convergencia entre los estudios, como las dificultades económicas, la historia personal y familiar de los estudiantes, las falencias en la educación media; y con ello plantear estrategias que puedan ser usadas por universidades. Con estos programas que son específicos para cada factor, se concluyó que, aunque existe gran interés en el fenómeno de la deserción y diversidad en los estudios encontrados, la insolvencia reposa en la separación de estrategias y políticas efectivas y cuestionables que frenen la deserción en las universidades.

CAPITULO III

MARCO METODOLÓGICO

En el presente estudio se utilizará la metodología Descubrimiento de Conocimiento en Bases de Datos o KDD para diseñar un modelo matemático que permita predecir qué variables socioeconómicas de los estudiantes influyen en la probabilidad de que los estudiantes abandonen sus estudios en la institución. Para determinar el modelo matemático, se utilizó una regresión logística multivariante en el programa R studio. Por último, se desarrolló un análisis estadístico para determinar la validez del modelo matemático propuesto.

3.1. Ubicación

Los datos serán obtenidos del Instituto Superior Tecnológico Luis A. Martínez Agronómico (ISTLAMA), ubicado en la ciudad de Ambato en la Avenida Rumiñahui y Pichincha esquina, como se muestra en la Figura 1.- Ubicación geográfica del Instituto Superior Tecnológico Luis A. Martínez Agronómico.



Figura 1. – Ubicación geográfica del Instituto Superior Tecnológico Luis A. Martínez Agronómico
Fuente: Google maps, 2020

3.2. Equipos y materiales

Para la realización de la presente investigación se utilizaron los siguientes materiales:

Hardware

- Laptop HP DESKTOP-MCVATE3 Intel (R) Core (TM) i7

- Impresora Epson L220

Software

- Base de datos de estudiantes
- Programa R Studio Versión 1.3.959 2019
- Paquete Microsoft Office 2016

3.3. Tipo de investigación

3.3.1. Investigación bibliográfica

La presente investigación será bibliográfica ya que serán utilizados libros, revistas, proyectos de titulación y artículos científicos para la construcción de la fundamentación teórica de la deserción estudiantil.

3.3.2. Investigación de campo

La investigación se considera también investigación de campo ya que se pretende obtener la información socioeconómica y su influencia en la tasa de deserción estudiantil de los estudiantes del Instituto Superior Tecnológico Luis A. Martínez Agronómico.

3.3.3. Investigación descriptiva

La investigación es de tipo descriptivo ya que todos los datos obtenidos serán descritos detalladamente, así como su influencia en los resultados finales.

3.4. Prueba de hipótesis

En la presente investigación, se consideran las siguientes hipótesis:

Hipótesis nula: El modelo matemático no es significativo ($\beta_1=\beta_2=\dots \beta_k=0$)

Hipótesis alternativa: El modelo matemático es significativo ($\beta_1 \neq \beta_2 \neq \dots \beta_k \neq 0$)

3.5. Población o muestra

La población de estudio corresponde a todos los estudiantes del Instituto Superior Tecnológico Luis A. Martínez Agronómico, instituto público de educación superior tecnológica de la ciudad de Ambato. Se estudiaron a 849 estudiantes de las carreras de Tecnología Superior en Procesamiento de Alimentos (296), Tecnología Superior en Gastronomía (299) y Tecnología Superior en Producción Pecuaria (254), matriculados en los períodos académicos Noviembre 2018-Abril 2019 y Noviembre 2019-Abril

2020. El Instituto Superior Tecnológico Luis A. Martínez Agronómico, ubicado en la ciudad de Ambato en la Avenida Rumiñahui y Pichincha esquina.

3.6. Recolección de información

Los datos recolectados para la presente investigación pertenecen a los estudiantes del Instituto Superior Luis A. Martínez Agronómico. La información se obtuvo de entrevistas realizadas a toda la población de estudiantes, como requisito de matriculación al inicio de cada período académico. La información se almacenó en una matriz en Microsoft Excel y se actualiza cada período académico. Cabe recalcar que como en el estudio de (Camacho et al., 2016) la presente investigación es de tipo estudio de caso, ya que estudia un sujeto o una realidad de carácter específico, a través de una investigación de campo realizada con información obtenida de los estudiantes y descriptiva, porque se describen las variables socioeconómicas de los estudiantes y su influencia en la probabilidad de desertar.

3.7. Procesamiento de la información y análisis estadístico

El procesamiento de datos se realizó siguiendo la metodología KDD, o generación de conocimiento en bases de datos, que es reconocido como un proceso no trivial para identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos (Fayyad, Piatetsky-Shapiro y Smyth, 1996). Las fases de la técnica se muestran en la Figura 2.- Metodología KDD.

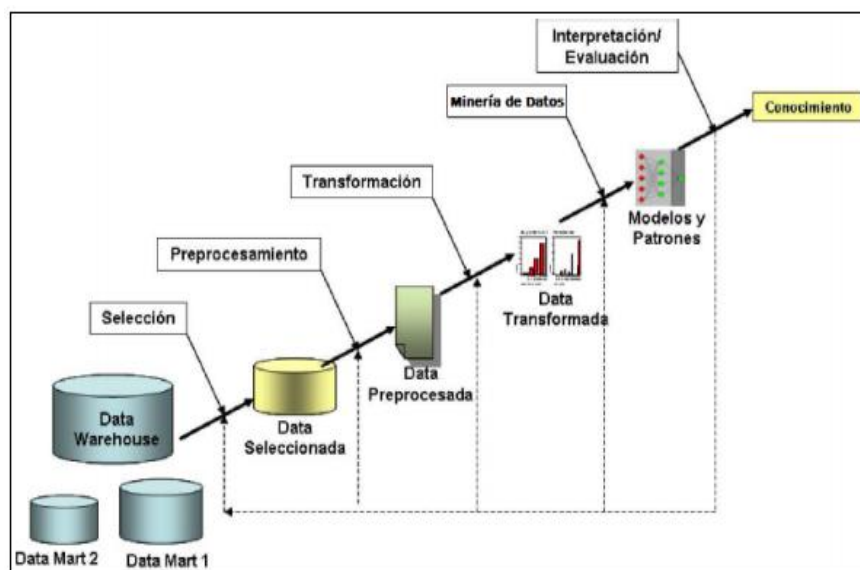


Figura 2. – Metodología KDD
Fuente: Fayyad, Piatetsky-Shapiro y Smyth, 1996)

Para el presente estudio se consideraron las siguientes variables independientes: género, estado civil, edad, carrera, repitencia, ocupación e ingresos económicos, como se indica en la Tabla 1.- Variables dependientes del proyecto de investigación. Las variables fueron seleccionadas con base a la revisión bibliográfica de estudios afines, sobre las variables socioeconómicas y académicas de los estudiantes que pueden influir en su posible deserción.

Tabla 1.- Variables dependientes del proyecto de investigación

Variables independientes	Descripción
Género	Femenino Masculino
Estado civil	Soltero Casado Unión libre Divorciado
Edad	Valor en años
Carrera	PA Procesamiento de Alimentos G Gastronomía PP Producción Pecuaria
Repitencia	Estudiante a reprobado al menos una asignatura Estudiante no ha reprobado ninguna asignatura
Ocupación	Estudiante estudia y trabaja Estudiante solamente estudia
Ingresos económicos	Valor promedio en USD de ingresos económicos percibidos mensualmente

Elaborado por: Vinueza, 2021

El 70% de los datos fue utilizado para generar el modelo predictivo requerido y el 30% restante para verificar su eficacia, como se indica en el trabajo de Zamorano y Martín (2018). El modelo predictivo se desarrolló mediante un Modelo Multivariante con Regresión Logarítmica, ya que los modelos de elección discreta son bastantes apropiados para analizar los factores determinantes de la probabilidad de un suceso. Adicionalmente se realizó una comparativa de datos con árboles de decisión.

3.8. Variables respuesta o resultados alcanzados

En la presente investigación se establecieron las siguientes variables

- **Variable Independiente:** Variables socioeconómicas de los estudiantes involucrados
- **Variable Dependiente:** Probabilidad de deserción de los estudiantes involucrados

Para evaluar la variable dependiente, se solicitaron datos a la secretaría de la institución, de las listas de matriculados de cada semestre, identificando cuales estudiantes desertaron. Para aplicar el modelo de regresión logística se adaptó la variable de respuesta a la retención del estudiante, siendo 0 si el estudiante no desertó y 1 si el estudiante desertó. Cabe recalcar que se consideró que un estudiante no desertó si se matriculó en el último período académico de la carrera académica correspondiente.

CAPITULO IV

RESULTADOS Y DISCUSIÓN

4.1. Análisis descriptivo de las variables socioeconómicas de los estudiantes involucrados en la investigación

4.1.1. Variables independientes cualitativas

En la Tabla 2.- Distribución porcentual en la muestra de las variables independientes se muestra la distribución porcentual de la población de las variables independientes cualitativas del estudio.

Tabla 2.- Distribución porcentual en la muestra de las variables independientes

Variable	Descripción	Total	Porcentaje
Género	Femenino	511	60,18 %
	Masculino	338	39,82 %
Estado civil	Soltero	762	89,75 %
	Divorciado	54	6,36 %
	Casado	11	1,30 %
	Unión libre	22	2,68 %
Carrera	Procesamiento de Alimentos	296	34,86 %
	Gastronomía	299	35,22 %
	Producción Pecuaria	254	29,92 %
Repitencia	Estudiante a reprobado al menos una asignatura	126	14,84 %
	Estudiante no ha reprobado ninguna asignatura	723	85,16 %
Ocupación	Estudiante estudia y trabaja	276	32,50 %
	Estudiante solamente estudia	573	67,50 %

Elaborado por: Vinuesa, 2021 (Fuente: IST Luis A. Martínez Agrónomico)

Con respecto al género, se encontró que el 60,18 % es de sexo femenino y el 39,82 % es de sexo masculino. Esta distribución se explica por la especialidad del ISTLAMA que resulta más atrayente para estudiantes del género femenino. En el estado civil de los estudiantes se observa que el 89,75 % son solteros, seguido del 6,36 % que es divorciado, el 2,68 % se encuentra en unión libre y el 1,30 % son casados. Adicionalmente, se observó un nivel de repitencia del 14,84% que corresponde a estudiantes que han repetido al menos una asignatura en la malla académica. Por último, se observa que el 32,50 % de los estudiantes indican que, adicionalmente a sus estudios de educación superior, trabajan; en el mayor de los casos para financiar sus estudios y mantener su hogar.

Las variables estudiadas se escogieron, debido a que la deserción por problemas personales puede verse afectada por el sexo del estudiante, afectando más a las mujeres que a los hombres, los conflictos de carácter económico y entorno familiar y social, se incluyen en esta causa. Se constató que los estudiantes de sexo masculino, vinculados al mercado laboral y provenientes de otras regiones del país, presentan mayor probabilidad de desertar (Baquerizo, Tam y López, 2014). Adicionalmente, según Sánchez (2016), en el Ecuador una de las principales razones que causan la deserción es la falta de acompañamiento a los jóvenes secundarios a la hora de escoger su carrera superior. Sánchez afirma que aproximadamente el 40% de los jóvenes no saben qué estudiar en la universidad, por lo cual los estudiantes optan por una carrera incorrecta, que conlleva finalmente a repitencia y deserción.

4.1.2. Variables independientes cuantitativas

De todas las variables de estudio ('género', 'estado civil', 'edad', 'carrera', 'repitencia', 'ocupación' e 'ingresos económicos'), las variables 'edad' e 'ingresos económicos' son las únicas cuantitativas. En la Figura 3. – Diagrama de caja y bigotes de la edad en función de la deserción estudiantil se observa el diagrama de caja y bigotes que relaciona 'edad' con 'deserción estudiantil'. La 'edad' presentó un valor promedio de 23,19 años, la edad mínima fue 18 años hasta 21 años valor del primer cuartil; posteriormente el cuartil dos, es decir, la mediana, presenta un valor de 22 años; por último, el tercer cuartil es 25 años hasta la edad máxima 59 años. Con respecto a las medidas de dispersión se observó que la varianza es 15,44, y la desviación estándar tiene un valor de 3,93.

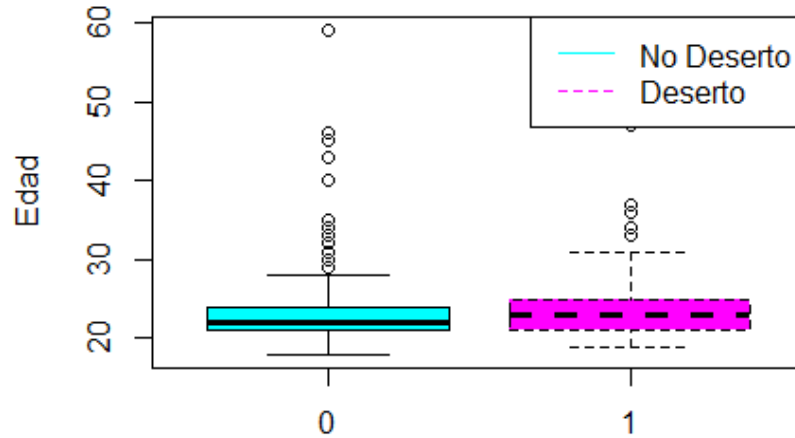


Figura 3. – Diagrama de caja y bigotes de la edad en función de la deserción estudiantil

Elaborado por: Vinueza, 2021

En la Figura 4. – Diagrama de caja y bigotes de la edad en función de la deserción estudiantil se muestra la variable ‘ingresos económicos’, la cual presentó un valor promedio de 560,21 USD, con un valor mínimo de 50 USD y máximo de 6 000 USD, que se considera un valor atípico. El primer cuartil corresponde a 360 USD, el segundo cuartil, es decir, la mediana fue 450 USD y el tercer cuartil corresponde 700 USD. La desviación estándar 438,68 USD.

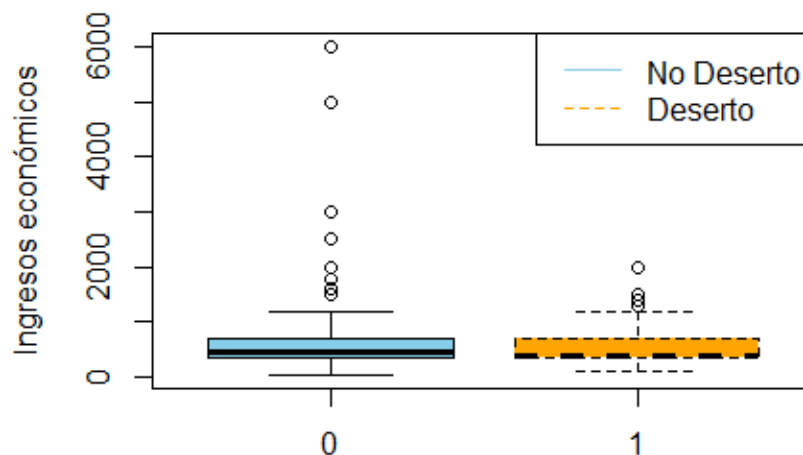


Figura 4. – Diagrama de caja y bigotes de la edad en función de la deserción estudiantil

Elaborado por: Vinueza, 2021

La importancia del estudio de los ingresos económicos de los estudiantes, se debe a que el nivel superior la deserción estudiantil se produce con mayor frecuencia en los primeros semestres académicos y el factor económico es el que mayor relevancia justifica estas deserciones (77,9 %); seguido de los problemas familiares (7,5%) y los relacionados con las insuficiencias en el desempeño del personal-académico (6,9%) (Yaselga y Yépez, 2010).

4.1.3. Variable dependiente

En la Figura 5.- Diagrama de barras de la deserción estudiantil del instituto, se muestra la

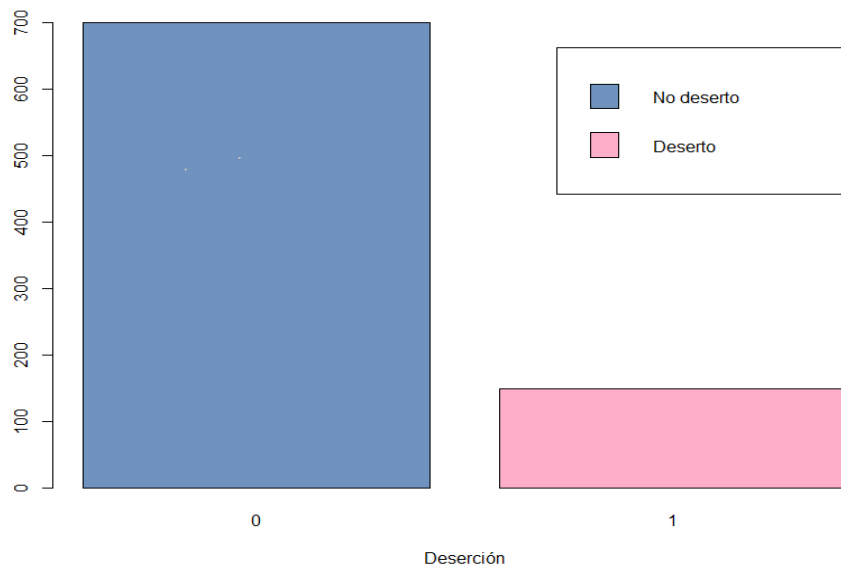


Figura 5. – Diagrama de barras de la deserción estudiantil del instituto
Elaborado por: Vinueza, 2021

La variable dependiente corresponde a la deserción de los estudiantes, se observó que el 82,45 % de estudiantes no desertaron y se matricularon en el último período académico de cada carrera, a diferencia del 17,55 % que abandonó la institución por varios motivos. En el estudio de Matallana, Gonzalez y Fonseca (2020) se realizó el seguimiento a 510 estudiantes en un periodo comprendido entre el 26 de marzo de 2019 y 26 de marzo de 2020, de los cuales 348 eran mujeres y 162 hombres, de los cuales 96 desertaron equivalente al 19%.

4.2. Aplicación de un modelo de análisis multivariante con regresión logarítmica para la obtención del modelo predictivo

Para realizar el análisis predictivo de los datos, se desarrolló un primer modelo de regresión logística con todas las variables independientes disponibles, obteniéndose los resultados de la Figura 6. – Resumen de modelo de regresión logística 1.

Coefficients:				
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.123e+00	9.692e-01	-1.159	0.2464
GéneroMasculino	-4.215e-02	2.312e-01	-0.182	0.8553
EstadocivilDivorciado	4.049e-01	8.175e-01	0.495	0.6204
EstadocivilSoltero	-3.167e-01	4.584e-01	-0.691	0.4896
EstadocivilUnión libre	-1.101e+00	1.140e+00	-0.966	0.3342
Repitenciasi	5.979e-01	2.731e-01	2.189	0.0286 *
OcupaciónTrabaja y estudia	3.974e-01	2.470e-01	1.609	0.1077
Ingresos	-8.573e-05	2.742e-04	-0.313	0.7545
Edad	-2.132e-03	2.977e-02	-0.072	0.9429
CarreraPA	-6.728e-01	2.742e-01	-2.454	0.0141 *
CarreraPP	-2.779e-01	2.707e-01	-1.027	0.3046

signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
(Dispersion parameter for binomial family taken to be 1)				
Null deviance: 547.57 on 592 degrees of freedom				
Residual deviance: 528.61 on 582 degrees of freedom				
(3 observations deleted due to missingness)				
AIC: 550.61				

Figura 6. – Resumen de modelo de regresión logística 1

Elaborado por: Vinueza, 2021

De los resultados obtenidos, se observó que solamente las variables ‘carrera’ (‘PA-Procesamiento de Alimentos’ y ‘PP-Producción Pecuaria’) y ‘repitencia’ resultaron significativas, con 95 % de confianza, ya que el valor de p es menor a 0,05. La evaluación del criterio de información de Akaike (AIC) del modelo arrojó un resultado de 550,61.

Para mejorar el modelo 1, se decidió proponer un segundo modelo, en el cual se eliminen las variables que presenten mayor valor de p, que en nuestro caso corresponden a las variables ‘Edad’ (p = 0,9429) y ‘Género’ (p = 0,8553). Los resultados del este nuevo modelo se presenta en la Figura 7. – Resumen de modelo de regresión logística 2.

En el estudio de Segura y Loza (2017) se reportó que las variables ‘beca académica’, ‘edad’, ‘provincia’ y ‘título de escuela secundaria’ influyen en el rendimiento académico de los estudiantes. Adicionalmente, González y Arismendi (2018) en su trabajo mostraron que los factores con mayor significancia estadística corresponden al ‘género’, ‘año de egreso de enseñanza media’ y ‘jornada de estudio’.

Coefficients:				
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.191e+00	4.623e-01	-2.575	0.0100 *
EstadocivilDivorciado	4.185e-01	8.141e-01	0.514	0.6072
EstadocivilSoltero	-3.085e-01	4.009e-01	-0.769	0.4416
EstadocivilUnión libre	-1.082e+00	1.124e+00	-0.963	0.3357
Repitenciasi	5.982e-01	2.730e-01	2.191	0.0284 *
OcupaciónTrabaja y estudia	3.847e-01	2.350e-01	1.637	0.1016
Ingresos	-9.285e-05	2.720e-04	-0.341	0.7329
CarreraPA	-6.718e-01	2.740e-01	-2.452	0.0142 *
CarreraPP	-2.772e-01	2.706e-01	-1.025	0.3056

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
(Dispersion parameter for binomial family taken to be 1)				
Null deviance: 547.57 on 592 degrees of freedom				
Residual deviance: 528.65 on 584 degrees of freedom				
(3 observations deleted due to missingness)				
AIC: 546.65				

Figura 7. – Resumen de modelo de regresión logística 2
Elaborado por: Vinueza, 2021

Así como en el modelo anterior, se observa que solamente las variables ‘carrera’ y ‘repitencia’ resultaron significativas, así como el intercepto, con 95 % de confianza, ya que el valor de p es menor a 0,05. El valor de AIC disminuye de 550,61 a 546,65, lo que significa que el modelo 2 es mejor que modelo 1.

Una vez más se eliminan nuevamente las variables que presentan mayor valor de p, es decir, ‘Estado civil’ (p = 0,6072) e ‘Ingresos económicos’ (p = 0,7329) y se obtiene el modelo 3, cuyo resultado se muestra en la Figura 8. – Resumen de modelo de regresión logística 3. El estudio de Albarrán (2019) revela que 65% de los estudiantes podrían dejar la institución por la falta de oportunidades laborales futuras, aumento de sus gastos personales y académicos, interrupción de las labores académicas por continuas protestas sociales, carencia de recursos económicos, baja formación escolar secundaria y desmotivación, principalmente.

```

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.5338    0.2098  -7.312 2.64e-13 ***
RepitenciaSi    0.5881    0.2710   2.170  0.0300 *
OcupaciónTrabaja y estudia  0.4270    0.2280   1.873  0.0611 .
CarreraPA     -0.6596    0.2715  -2.430  0.0151 *
CarreraPP     -0.3040    0.2683  -1.133  0.2572
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 547.57  on 592  degrees of freedom
Residual deviance: 530.82  on 588  degrees of freedom
(3 observations deleted due to missingness)
AIC: 540.82

```

Figura 8. – Resumen de modelo de regresión logística 3
Elaborado por: Vinueza, 2021

En el modelo 3, se mantienen las variables ‘carrera’ y ‘repitencia’ como variables significativas, así como el intercepto, con 95 % de confianza, ya que el valor de p es menor a 0,05. El valor de AIC disminuye de 546,65 a 540,82, lo que indica que el modelo está mejorando. Por último, se elimina la variable ‘ocupación’ (p = 0,611), obteniéndose el modelo 4, mostrado en la Figura 9. – Resumen de modelo de regresión logística 4.

```

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.3305    0.1745  -7.622 2.49e-14 ***
RepitenciaSi  0.6068    0.2698   2.249  0.02451 *
CarreraPA    -0.7494    0.2664  -2.813  0.00492 **
CarreraPP    -0.3752    0.2649  -1.417  0.15656
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 547.57  on 592  degrees of freedom
Residual deviance: 534.27  on 589  degrees of freedom
(3 observations deleted due to missingness)
AIC: 542.27

```

Figura 9. – Resumen de modelo de regresión logística 4
Elaborado por: Vinueza, 2021

El modelo 4 se determinó como el mejor modelo predictivo, ya que las variables ‘repitencia’ y ‘carrera’ son significativas, con base a los valores de sus coeficientes se lo puede representar en la siguiente fórmula:

$$P(Y) = \frac{1}{1 + e^{(1.33 - 0.61x_1 + 0.75x_2 + 0.38x_3)}} \quad \text{Ecuación 1}$$

En el modelo 4 se observa que las dos variables ‘repitencia’ y ‘carrera’ (Procesamiento de alimentos) son significativas, con un valor p de 0,0245 y 0,0049 respectivamente, presentando ambas variables valores de p menores a 0,05. El intercepto también se muestra como un valor significativo con un valor de p igual a $2,49 \times 10^{-14}$.

El diseño del modelo matemático obtenido se basa en la identificación de las variables académicas significativas (‘repitencia’ y ‘carrera’) que influyen directamente en la deserción de los estudiantes. Para el diseño del modelo se aplicó la herramienta R Studio, mediante la cual se obtuvo la sintaxis del programa que puede ser utilizado con una nueva base de datos y establece el modelo que se representa en la Ecuación 1.

Cuando se compara el AIC del modelo 3 (540,82) con el valor AIC del modelo 4 (542,27), se observa un aumento del valor, es decir, según el criterio Akaike, el modelo 3 es mejor que el modelo 4. Sin embargo, el criterio AIC si disminuyó si se compara el modelo 1 (550,60) con el modelo 4 (542,27), por lo tanto, se escoge el modelo 4 que puede explicarse con solamente dos variables, ‘repitencia’ y ‘carrera del estudiante’. El modelo 4 posteriormente se validó con otros métodos.

En el programa se determinó el valor de p es 0,000121 el cual es menor a 0,05; por lo tanto, el modelo es significativo con 95 % de confianza, como se muestra en la Figura 10. – Análisis de significancia del modelo de regresión logística 4. Esto permite que se rechace la hipótesis nula.

```
> #Estadístico de prueba
> with(mod4,null.deviance-deviance)
[1] 13.29436
>
> #Valor de P del estadístico de prueba
> with(mod4,pchisq(null.deviance-deviance,df.null-df.residual,lower.tail = FALSE))
[1] 0.004041405
> #El valor de p es 0.000121 menor a 0.05 por lo tanto el modelo es significativo con
> #95 % de confianza
> |
```

Figura 10. – Análisis de significancia del modelo de regresión logística 4
Elaborado por: Vinueza, 2021

Posteriormente, se realizó el análisis ANOVA del modelo 4, el cual se muestra en la Figura 11. – Análisis ANOVA del modelo de regresión logística 4.

```

> anova(mod4, test = "Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: Deserción

Terms added sequentially (first to last)

      Df Deviance Resid. Df Resid. Dev Pr(>Chi)
NULL                    592     547.57
Repitencia  1      5.0204     591     542.55 0.02505 *
Carrera     2      8.2740     589     534.27 0.01597 *

```

Figura 11. – Análisis ANOVA del modelo de regresión logística 4
Elaborado por: Vinueza, 2021

Del análisis ANOVA se observa que el valor p de ‘repitencia’ y ‘carrera’ es menor a 0,05 y por tanto significativo con un nivel de confianza de 95 %.

Posteriormente se determinan los coeficientes y razón Odd del modelo 4, como se muestra en la Tabla 3. - Coeficientes y valores de razón Odd del modelo de regresión logística 4.

Tabla 3.- Coeficientes y valores de razón Odd del modelo de regresión logística 4

Variables significativas	Coeficientes	Razón Odd
Repitencia (Si)	0,5539975	1,7401955
Carrera (Procesamiento de Alimentos)	-0,8537485	0,4258158
Carrera (Producción Pecuaria)	-1,0143231	0,3626478

Elaborado por: Vinueza, 2021
Fuente: Programa R Studio, 2021

Se observa que para la variable ‘repitencia’, la razón Odd es 1,74. Por lo tanto, se puede afirmar que el hecho de que un estudiante repita alguna asignatura aumenta en un 74 % la probabilidad de que abandone la institución. Por otro lado, con respecto a la variable ‘carrera’, el valor de la razón Odd de la carrera ‘Procesamiento de

Alimentos' es 0,42, lo que muestra que los estudiantes de esta carrera tienen 58% menor probabilidad de desertar.

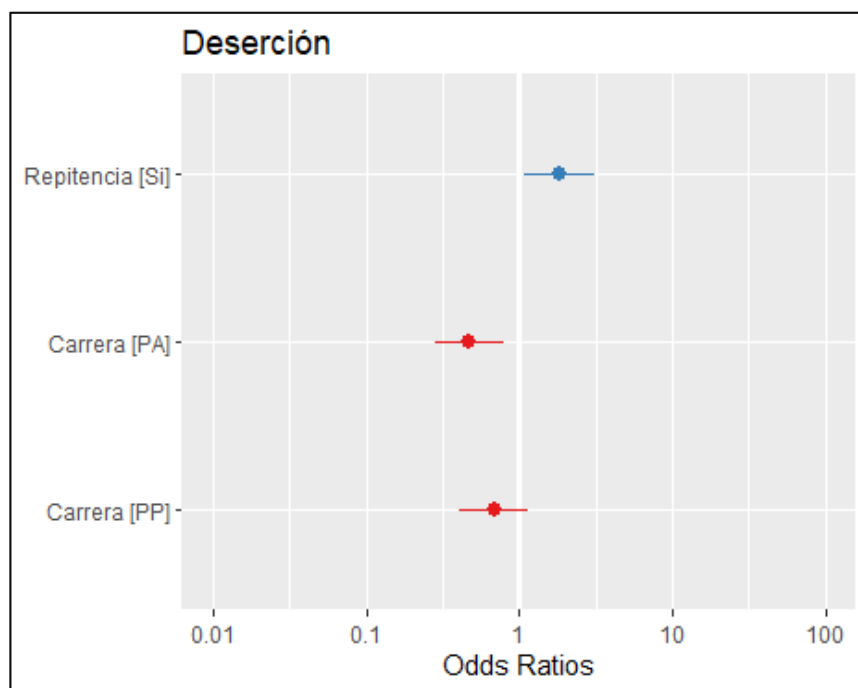


Figura 12. – Valores de razón Odd de las variables repitencia y carrera del modelo de regresión logística 4
Elaborado por: Vinueza, 2021

Por su lado, la razón Odd de la carrera de 'Producción Pecuaria' es 0,36; lo que permite concluir que los estudiantes de esta carrera tienen 64% menor probabilidad de desertar. Esto se puede observar en la Figura 12. – Valores de razón Odd de las variables repitencia y carrera del modelo de regresión logística 4.

En el estudio de Barahona, Veres y Aliaga (2016) quienes, utilizaron un modelo de regresión logística para determinar aquellos factores que han incidido en la tasa de deserción, el valor de la razón Odd, indica que los créditos (asignaturas) inscritos reducen en un 9 % la probabilidad de desertar. Es decir, los estudiantes que no desertaron son aquellos con la mayor cantidad de créditos inscritos y que han realizado un mayor esfuerzo académico. De igual forma sucede con la variable rendimiento académico que disminuye en 14 % veces la probabilidad de desertar

Adicionalmente, en la Tabla 4- Comparación de cuatro modelos de regresión logística estudiados se muestra una comparación entre los cuatro modelos establecidos, donde se muestra la significancia de las variables independientes establecidas.

Tabla 4.- Comparación de cuatro modelos de regresión logística estudiados

Variable independiente	Variable dependiente			
	Deserción			
	(1)	(2)	(3)	(4)
Género: masculino	0,959 (1,260)			
Estado civil: divorciado	1,499 (2,265)	1,520 (2,257)		
Estado civil: soltero	0,729 (1,581)	0,735 (1,493)		
Estado civil: Unión libre	0,333 (3,126)	0,339 (3,076)		
Repitencia: Si	1,818 (1,314)	1,819 (1,314)	1,801 (1,311)	1,835 (1,310)
Ocupación: trabaja y estudia	1,488 (1,280)	1,469 (1,265)	1,533 (1,256)	
Ingresos	1,000 (1,000)	1,000 (1,000)		
Edad	0,998 (1,030)			
Carrera: Procesamiento de Alimentos	0.510 (1.315)	0.511 (1.315)	0.517 (1.312)	0.473 (1.305)
Carrera: Producción Pecuaria	0.757 (1.311)	0.758 (1.311)	0.738 (1.308)	0.687 (1.303)
Constante	0.325 (2.636)	0.304 (1.588)	0.216 (1.233)	0.264 (1.191)
Probabilidad logarítmica	-264,304	-264,324	-265,410	-267,137
Criterio Akaike	550,609	546,648	540,820	542,275

Elaborado por: Vinueza, 2021

Fuente: Programa R Studio, 2021

A continuación, se realizó la validación del modelo utilizando una Prueba Razón de verosimilitud, comparando el modelo 1, que implicaba todas las variables independientes, con el modelo 4 que considera solamente las variables ‘carrera’ y ‘repetencia’. Realizada la prueba, se obtiene un valor de p igual a 0,5793, mayor a 0,05. Por lo tanto, el modelo 1 es estadísticamente igual a modelo 4. El resultado del test se muestra en la Figura 13. – Prueba de razón de similitud entre modelo 1 y modelo 4.

Likelihood ratio test					
Model 1: Deserción ~ Género + Estadocivil + Repitencia + Ocupación + Ingresos + Edad + Carrera					
Model 2: Deserción ~ Repitencia + Carrera					
#Df	LogLik	Df	Chisq	Pr(>Chisq)	
1	11	-264.30			
2	4	-267.14	-7	5.6659	0.5793

Figura 13. – Prueba de razón de similitud entre modelo 1 y modelo 4
Elaborado por: Vinueza, 2021

Posteriormente, con respecto a los datos de entrenamiento (70%), se obtuvo que el porcentaje de datos que se clasifican correctamente es del 83 % según la matriz de confusión que muestra en la Tabla 5.

Tabla 5.- Matriz de confusión de datos de entrenamiento

		PREDICCIÓN	
		No desertó	Desertó
REAL	No desertó	497	1
	Desertó	96	0

Elaborado por: Vinueza, 2021

Fuente: Programa R Studio, 2021

En la Tabla 5.- Matriz de confusión de datos de entrenamiento, se puede observar que 497 de los datos fueron clasificados correctamente, mientras 97 fueron clasificados incorrectamente.

4.3. Implementación el modelo predictivo para determinación de probabilidad de deserción de los estudiantes involucrados.

Para la implementación del modelo predictivo se consideró el 30 % de datos que no fueron utilizados en el entrenamiento. En las pruebas, se encontró que el 79 % de los

datos de testeo fueron clasificados correctamente. La matriz de confusión de los datos de testeo se muestra en la Tabla 6.- Matriz de confusión de datos de testeo.

Tabla 6.- Matriz de confusión de datos de testeo

		PREDICCIÓN	
		No desertó	Desertó
REAL	No desertó	202	0
	Desertó	53	0

Elaborado por: Vinueza, 2021

Fuente: Programa R Studio,2021

En la Tabla 6.- Matriz de confusión de datos de testeo., se observa que 202 de los datos fueron clasificados correctamente, mientras que 53 fueron clasificados incorrectamente. Posteriormente se calcularon los valores de precisión, exhaustividad (recall) y F1-score de la matriz de confusión de los datos de testeo. Los resultados obtenidos se muestran en la Tabla 7.- Valores de evaluación de la matriz de confusión de datos de testeo.

Tabla 7.- Valores de evaluación de la matriz de confusión de datos de testeo

Parámetros	Valores
Precisión	1,00
Exhaustividad (Recall)	0,79
F1-Score	0,88

Elaborado por: Vinueza, 2021

Fuente: Programa R Studio,2021

Los valores del F1-score muestran que el modelo matemático número 4, presenta un rendimiento del 88% en la clasificación de estudiantes que desertan de la institución.

Por motivos de comparación, se desarrolló un modelo de predicción con la técnica de Árbol de decisión, los resultados se muestran en la Figura 14. – Árbol de decisión podado a dos niveles.

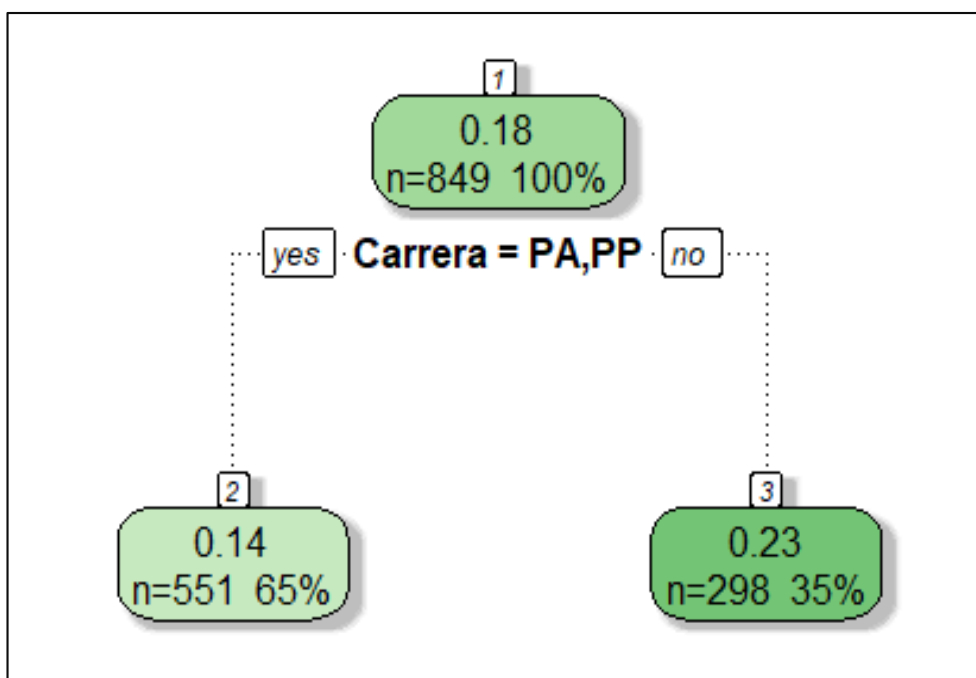


Figura 14. – Árbol de decisión podado a dos niveles
Elaborado por: Vinueza, 2021

El Árbol de decisión determinó que la variable ‘carrera’ es la única variable necesaria para predecir la deserción estudiantil dentro de la institución, a diferencia de modelo de regresión logística número 4 que concluyó que las variables significativas fueron ‘repitencia’ y ‘carrera’. La matriz de confusión de los datos predichos con el árbol de decisión se muestra en la Tabla 8. - Matriz de confusión de árbol de decisión.

Tabla 8.- Matriz de confusión de árbol de decisión

		PREDICCIÓN	
		No desertó	Desertó
REAL	No desertó	472	228
	Desertó	79	70

Elaborado por: Vinueza, 2021

Fuente: Programa R Studio,2021

Los valores de precisión, exhaustividad (recall) y F1-score de la matriz de confusión del árbol de decisión se presentan en la Tabla 9.- Valores de evaluación de la matriz de confusión de árbol de decisión.

Tabla 9.- Valores de evaluación de la matriz de confusión de árbol de decisión

Parámetros	Valores
Precisión	0,67
Exhaustividad (Recall)	0,86
F1-Score	0,74

Elaborado por: Vinueza, 2021

Fuente: Programa R Studio,2021

Se encontró que el valor F1-Score del modelo de regresión logística 4 (0,88) es mayor a valor F1-Score obtenido con el árbol de decisión (0,74), lo que puede deberse al número de variables explicativas de cada modelo, dos en la regresión logística y una en el árbol de decisión.

Jurado y Plata (2019) en su investigación, para diseñar un modelo predictivo considerando cuatro modelos de clasificación: Potenciación del Gradiente Estocástico (GBM), Bosques Aleatorios (RF), Redes Neuronales (NNET) y Regresión Logística (GLM); además, se utilizó la metodología KDD. Se concluyó que los modelos Bosques Aleatorios y Redes Neuronales se destacaron más, pero se escogió el modelo Bosques Aleatorios por su facilidad de uso.

Los resultados del presente estudio, que concluyen que la variable 'repitencia' y 'carrera', concuerdan con los obtenidos en la investigación de Vicente (2020) quien encontró que las variables calificaciones de los niveles de segundo, cuarto y quinto semestre fueron las más influyentes en el modelo de predicción de deserción estudiantil. El presente estudio y con el trabajo de Cuji, Gavilanes y Sánchez (2017), con la metodología de árbol de decisión quienes determinaron que las variables nivel y notas tienen mayor influencia en la deserción. Cabe recalcar que repitencia de se encuentra directamente relacionada con las calificaciones y rendimiento académico de los estudiantes.

Por el contrario, y con respecto al estudio de Sopalo, Guevara y Burbano (2020), quien indica que pertenecer al género femenino y vivir en el área rural aumentan la probabilidad de deserción, en el presente trabajo la variable género no se considera significativa. De igual forma Segura y Loza (2017) encontraron que una de las variables que influyen en el rendimiento académico, es la edad, sin embargo, para el caso de estudio descrito, ésta no se presentó como una variable significativa. En resumen, se identificaron las variables ‘repitencia’ y ‘carrera’ como las únicas variables significativas, si los estudiantes repiten alguna asignatura tienen 74 % mayor probabilidad de desertar. Con respecto a la carrera, los estudiantes que estudian en ‘Procesamiento de Alimentos’ y ‘Producción Pecuaria’ tienen 58% y 64 % menor probabilidad de desertar respectivamente. Adicionalmente se encontró que el 83 % de los datos de entrenamiento fueron clasificados correctamente, que difiere levemente de los datos de testeos que fueron clasificados correctamente en un 79%.

4.4. Verificación de hipótesis

1. Planteamiento de hipótesis

1. Modelo lógico

Hipótesis nula: El modelo matemático no es significativo

Hipótesis alternativa: El modelo matemático es significativo

2. Modelo matemático

- $H_0: \beta_1 = \beta_2 = \dots \beta_k = 0$
- $H_1: \beta_1 \vee \beta_2 \vee \dots \beta_k \neq 0$

3. Resultados

El diseño del modelo predictivo muestra que los valores de los coeficientes son:

- $\beta_0 = -1.33$
- $\beta_1 = 0.60$
- $\beta_2 = -0.74$

- $\beta_0 = -0.37$

Por lo tanto, se aceptó la hipótesis alternativa que dice: “El modelo matemático es significativo ($\beta_1 \neq \beta_2 \dots \beta_k \neq 0$)”, ya que cuatro de los coeficientes del modelo son diferentes de cero, como se muestra en la Ecuación 1. Por lo tanto, se rechaza la hipótesis nula, que dice: “El modelo matemático no es significativo ($\beta_1 = \beta_2 = \dots \beta_k = 0$)”.

CAPITULO V

CONCLUSIONES, RECOMENDACIONES, BIBLIOGRAFÍA Y ANEXOS

5.1. Conclusiones

- Se realizó el análisis descriptivo de las variables de estudio, los resultados más representativos indican que 60, 18 % de los estudiantes son de género femenino, el 89,75 % son solteros, el 32,50 % trabajan y estudian y se encontró un nivel de repitencia del 14,84 %. La edad promedio de los estudiantes es 23,19 años y el valor promedio de ingresos económicos es 560, 21 USD.
- Se aplicó un modelo de análisis multivariante con regresión logarítmica para obtener un modelo predictivo, concluyó que las variables significativas son ‘repitencia’ y ‘carrera’ a la que pertenece el estudiante resultaron ser las únicas significativas.
- Se encontró que la variable significativa ‘repitencia’ indica que los estudiantes asocian el fracaso académico con su deseo de deserción. Según la razón Odd de la variable ‘repitencia’ si un estudiante repite alguna asignatura aumenta en un 74 % la probabilidad de que abandone la institución.
- Se encontró que la variable significativa ‘carrera’ indica que los estudiantes de la Carrera de ‘Gastronomía’ tiene mayor probabilidad de desertar que de las otras dos carreras (‘Procesamiento de Alimentos’ y ‘Producción Pecuaria’). En la variable ‘carrera’, el valor de la razón Odd de ‘Procesamiento de Alimentos’ es 0,42 y 0,36 de carrera de ‘Producción Pecuaria’, es decir, los estudiantes tienen 58% y 68 % menor probabilidad de desertar, respectivamente.
- Se implementó el modelo predictivo para estimar deserción estudiantil, y se observó que el modelo de regresión logística 4 clasificó correctamente el 83 % de los datos de entrenamiento y el 79 % de los datos de testeo.
- El valor F1_Score del modelo de regresión logística 4 fue mayor que el valor del F1_Score del modelo con árbol de decisión, concluyéndose que el rendimiento del modelo de regresión logística (88%) es mejor que el modelo con árbol de decisión (74%).

5.2. Recomendaciones

- La institución debe fortalecer las actividades de seguimiento académico por parte de los docentes y Bienestar Institucional, para evitar la repetencia de asignaturas de los estudiantes, y, por tanto, su deseo de desertar.
- La variable significativa ‘carrera’, en específico Carrera de Gastronomía requiere que la institución desarrolle una investigación más profunda sobre los posibles factores que inducen a los estudiantes de la carrera a desertar en mayor proporción que los estudiantes de las otras carreras.
- Las instituciones de educación superior deben considerar utilizar el modelo de predicción que utiliza el análisis multivariante con regresión logarítmica, para estimar la deserción estudiantil, ya que presenta mayor rendimiento que el modelo con árboles de decisión.
- En la presente investigación se recomienda establecer un modelo de regresión logística y modelo de árbol de decisión, por carrera, para verificar si existen otras variables significativas en el modelo.
- Indagar a los estudiantes sobre nuevas variables que puedan influir en la deserción para considerarlas en nuevos modelos de regresión logística, mediante encuestas y cuestionarios.
- Aplicar el modelo de regresión logística generado, en otras instituciones de educación superior de la provincia y zona centro del país.

5.3. BIBLIOGRAFÍA

- Albarrán-Peña, J. (2019). La deserción estudiantil en la Universidad de Los Andes (Venezuela). *Educación y Humanismo*, 21(36), 60-92.
- Barahona, P., Veres, E y Aliaga, V. (2016). Deserción académica de la Universidad de Atacama, Chile. *Comunicación*, 7(2), 27-37.
- Baquerizo, P., Tam, A. y López, J. (2014). La deserción y la repitencia en las instituciones de Educación Superior: algunas experiencias investigativas en el Ecuador. *Universidad y Sociedad*, 6(1).
- Camacho, M., Bravo, M., Bravo, M., y López, G. (2016). *Abandono estudiantil en la educación superior: Caso de estudio sobre las matemáticas básicas como factor asociado*. Congreso CLABES.
- Camacho, M., Montalvo, A y Galezo Arango, P. (2019). Determinantes de la deserción estudiantil en estudiantes universitarios.
- Campoverde, S. (2020). Deserción Estudiantil en el Instituto Tecnológico Superior de Fútbol, Guayaquil, 2019. (Tesis de grado). Universidad César Vallejo, Piura.
- Castro, J. (2020). *Minería de datos educacionales para determinar de forma temprana la deserción estudiantil*. (Tesis de grado). UTE, Quito.
- CEDIA. (2020). *Indicadores para la gestión de la calidad en la educación Superior Ecuatoriana*. Recuperado de <https://www.cedia.edu.ec/dmdocuments/publicaciones/Libros/indicadores2020.pdf>
- Cuji, B., Gavilanes, W. y Sánchez, R. (2017). Modelo predictivo de deserción estudiantil basado en arboles de decisión. *Espacios*, 38, 17-26.
- Fayyad, U., Piatetsky-Shapiro, G. y Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27-34.
- Frutos, C y Bueñano, E. (2017). *Análisis de la tasa de retención y su incidencia en la detección de patrones de deserción estudiantil en la Universidad Técnica de Ambato*. (Tesis de posgrado). Universidad Técnica de Ambato, Ambato.
- González, F y Arismendi, K. (2018). Deserción estudiantil en la educación superior técnico-profesional: Explorando los factores que indican en alumnos del primer año. *Revista de la Educación Superior*, 47(188), 109-137.

- Jurado, M y Plata, W. (2019). *Diseño de un modelo predictivo de la deserción estudiantil de postgrado en una institución de educación superior*. (Tesis de grado). Escuela Superior Politécnica de Litoral, Guayaquil.
- Himmel, E. (2002). Modelo de análisis de la deserción estudiantil en la educación superior. *Calidad en la Educación* (17),91-108. doi:<https://doi.org/10.31619/caledu.n17.409>.
- Machado, J y Castillo, L. (2019). *Modelado del tiempo de estancia estudiantil en la FICFM de la Universidad Central del Ecuador*. (Tesis de posgrado). Universidad Central del Ecuador, Quito.
- Matallana, A., González, J. y Fonseca, L. (2020). Modelo sobrevida para la deserción estudiantil de los programas de nivel técnico en una IES de formación para el trabajo en la ciudad de Bogotá.
- Montenegro, C., y Taco, M. (2012). *Repitencia y deserción de los estudiantes de Arquitectura y Administración: Carrera de Administración pública presencial y semi-presencial de la Universidad Central del Ecuador, causas, consecuencias y costos en los años lectivos 2004-2005 al 2008-2009*. (Tesis de Licenciatura). Quito- Ecuador.
- Pérez-Gutiérrez, B. R. (2020). Comparación de técnicas de minería de datos para identificar indicios de deserción estudiantil, a partir del desempeño académico. *Revista UIS Ingenierías*, 19(1), 193-204.
- Pérez, M., Escobar, R., Toledo, R., Gutierrez, B y Reyes, M. (2018). Modelo de predicción de la deserción estudiantil de primer año en la Universidad Bernardo O´ Higgins. *Educação e Pesquisa: Revista da Faculdade de Educação da Universidade de São Paulo*, 44, 86.
- Poveda, J. Poveda, I y España, I. (2020). Análisis de la deserción estudiantil en una universidad pública de Bolivia. *Revista Iberoamericana*, 82(2), 151-172.
- Reyes, J, Rosero, S y Medina, N. (2019). *Deserción escolar en estudiantes de 5 a 17 años en las Provincias de Pichincha y Guayas en el período 2013-2017*. (Tesis de grado). Universidad Central del Ecuador, Quito.
- Rodríguez, J. (2016). *Determinantes de la deserción escolar en la educación secundaria asociados a las características de las instituciones educativas en el Ecuador*. (Tesis de grado). Pontificia Universidad Católica del Ecuador, Quito.

- Sánchez, T. (2016). *La deserción universitaria*. Recuperado de <https://www.eltelegrafo.com.ec/noticias/sociedad/4/la-desercion-universitaria-bordea-el-40>.
- Segura-Morales, M., & Loza-Aguirre, E. (2017). Using Decision Trees for Predicting Academic Performance Based on Socio-Economic Factors. In *2017 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 1132-1136). IEEE.
- Sopalo, S., Guevara, G y Burbano, R. (2020). *Análisis de los factores que inciden en la deserción estudiantil de los niños, niñas y adolescentes ecuatorianos en el periodo 2009 – 2019*. (Tesis de grado). Escuela Politécnica Nacional, Quito.
- Tapasco, O., Ruiz, J., Osorio, D. y Ramírez, D. (2019). Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión. *Educación y Educadores*, 22(1), 81-100.
- Vicente, V y Cuji, B. (2020). *Aplicación de la técnica de minería de datos para la predicción de la deserción estudiantil universitaria*. (Tesis de grado). Universidad Técnica de Ambato, Ambato.
- Vizúete, R y Chasiluisa, I. (2017). *Deserción Escolar de los estudiantes del colegio técnico Vicente Rocafuerte de la ciudad de Quito de la especialidad de mecánica automotriz y la repercusión en el ámbito laboral (año 2017-2017)*. Universidad Central del Ecuador, Quito.
- Yaselga, B., y Yépez, P. (2010). *Factores que intervienen en la deserción de los estudiantes de segundo y cuarto semestres de la Carrera de Enfermería, Facultad Ciencias de la Salud*. Universidad Técnica del Norte (Tesis de Licenciatura). Ibarra- Ecuador.
- Yin, R.K. (2003). *Case study research: Design and Methods*. Recuperado de https://iwansuharyanto.files.wordpress.com/2013/04/robert_k-_yin_case_study_research_design_and_mebookfi-org.pdf
- Zambrano, G. Rodríguez, K y Guevara, L. (2018). Análisis de la deserción estudiantil en las universidades del Ecuador y América Latina. *Revista Pertinencia Académica*, 8(1), 1-28.
- Zamorano, J. (2018). Comparativa y análisis de algoritmos de aprendizaje automático para la predicción del tipo predominante de cubierta arbórea.

5.4. ANEXOS

Anexo 1.- Carta de aceptación del IST Luis A. Martínez Agronómico para uso de datos socioeconómicos y académicos de estudiantes involucrados



INSTITUTO SUPERIOR TECNOLÓGICO
LUIS A. MARTÍNEZ
(Agronómico)

CARTA DE ACEPTACIÓN

LA SUSCRITA RECTORA DEL INSTITUTO SUPERIOR TECNOLÓGICO
LUIS A. MARTÍNEZ AGRONÓMICO, TIENE A BIEN

CERTIFICAR

Que la Ingeniera Cristina Nataly Vinuesa López, quien presta sus servicios profesionales en calidad de Docente en el nivel superior en la institución, ha solicitado se le permita utilizar los datos del IST Luis A. Martínez Agronómico para realizar el proyecto de titulación “Implementación de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica”, para la obtención del título de Magister en Matemática Aplicada en la Universidad Técnica de Ambato. Requerimiento que ha sido **ACEPTADO**, considerando que los datos proporcionados sean utilizados solamente con fines académicos y con estrictos criterios de confidencialidad, sin que puedan causar daño a la imagen de ningún estudiante involucrado ni del IST Luis A. Martínez Agronómico.

Ambato, 09 de septiembre de 2019

Atentamente,


Lic. Maryuxi Castillo Mg
RECTORA ITSLAMA



Av. Rumiñahui y Pichincha Alta
Teléfono 032 419293
itslamagronomico@gmail.com

Anexo 2.- Cuestionario aplicado a los estudiantes para obtener la base de datos



**INSTITUTO SUPERIOR TECNOLÓGICO
LUIS A. MARTÍNEZ AGRONÓMICO**
FORMULARIO DE MATRÍCULA

SECRETARÍA DE
EDUCACIÓN SUPERIOR, CIENCIA,
TECNOLOGÍA E INNOVACIÓN



1. Tipo de Documento de Identificación: <input type="checkbox"/> Cédula <input type="checkbox"/> Pasaporte		Formulario Nro. _____
2. Número de Documento de Identificación: <input type="text"/>	3. Sexo: <input type="checkbox"/> Hombre <input type="checkbox"/> Mujer	4. Género: <input type="checkbox"/> Masculino <input type="checkbox"/> Femenino
5. Apellidos y Nombres del Estudiante:		
6. Correo Electrónico:		7. Nro. Celular:
9. Dirección:		8. Nro. Convencional:
11. En caso de emergencia contactar a: (Apellidos y Nombres)		10. Código postal:
12. Parentesco:		13. Nro. de Contacto:
14. Etnia: <input type="checkbox"/> Indígena <input type="checkbox"/> Afroecuatoriano <input type="checkbox"/> Negro <input type="checkbox"/> Mulato <input type="checkbox"/> Montuvio <input type="checkbox"/> Mestizo <input type="checkbox"/> Blanco <input type="checkbox"/> Otro <input type="checkbox"/> No registra <input type="checkbox"/>		
15. En caso de escoger etnia <i>Indígena</i> u <i>Otro</i> especifique su nacionalidad:		
16. Habla algún idioma ancestral: <input type="checkbox"/> Si <input type="checkbox"/> No Especifique:		
17. Fecha de Nacimiento: (Año-Mes-Día) <input type="text"/>		17.1 Edad: <input type="text"/>
18. Tipo de Sangre:		
19. País de Nacionalidad:	20. Provincia de Nacimiento:	21. Cantón de Nacimiento:
22. Categoría migratoria: (en caso de ser extranjero) <input type="checkbox"/> Residente permanente <input type="checkbox"/> Residente temporal <input type="checkbox"/> Residente transitorio o no residente <input type="checkbox"/> Refugiado		
23. País de Residencia:	24. Provincia de Residencia:	25. Cantón de Residencia:
26. Estado Civil: <input type="checkbox"/> Soltero <input type="checkbox"/> Casado <input type="checkbox"/> Unión Libre <input type="checkbox"/> Divorciado <input type="checkbox"/> Viudo		
27. Tiene Alguna Discapacidad: <input type="checkbox"/> Si <input type="checkbox"/> No		28. Nro. Carnet CONADIS:
29. Porcentaje de Discapacidad <input type="text"/> %	30. Tipo: <input type="checkbox"/> Auditiva <input type="checkbox"/> Física <input type="checkbox"/> Visual <input type="checkbox"/> Otra <input type="checkbox"/> Intelectual <input type="checkbox"/> Mental <input type="checkbox"/> NA	
31. Tipo de Colegio: <input type="checkbox"/> Fiscal <input type="checkbox"/> Fiscomisional <input type="checkbox"/> Particular <input type="checkbox"/> Municipal <input type="checkbox"/> Extranjero <input type="checkbox"/> No Registra		
32. Tipo de Bachillerato: <input type="checkbox"/> Técnico <input type="checkbox"/> Técnico Productivo <input type="checkbox"/> BGU <input type="checkbox"/> BI <input type="checkbox"/> Otro		33. Año de graduación:
34. Posee algún título de Educación Superior: <input type="checkbox"/> Si <input type="checkbox"/> No Especifique el título:		
35. Fecha en la que inició el estudiante la Carrera: (Año-Mes-Día) <input type="text"/>		36. Fecha de Matrícula: (Año-Mes-Día) <input type="text"/>
37. Tipo de Matrícula: <input type="checkbox"/> Ordinaria <input type="checkbox"/> Extraordinaria <input type="checkbox"/> Especial		
38. Período /Ciclo Académico: <input type="checkbox"/> Primero (Mayo-Oct) <input type="checkbox"/> Segundo (Nov- Abril)		39. Año del Período Académico: <input type="text"/>
40. Nivel Académico: <input type="checkbox"/> 1º <input type="checkbox"/> 2º <input type="checkbox"/> 3º <input type="checkbox"/> 4º <input type="checkbox"/> 5º <input type="checkbox"/> 6º		41. Paralelo (A-T):
42. Nombre de la Carrera:		43. Título que otorga la Carrera:
44. Tipo de Carrera: <input type="checkbox"/> Tecnicatura <input type="checkbox"/> Tecnología		45. Modalidad de la Carrera: <input type="checkbox"/> Presencial <input type="checkbox"/> Semipresencial <input type="checkbox"/> Dual
46. Jornada en la que Estudia: <input type="checkbox"/> Matutina <input type="checkbox"/> Vespertina <input type="checkbox"/> Nocturna <input type="checkbox"/> Intensiva		
47. Ha repetido al menos una materia: <input type="checkbox"/> Si <input type="checkbox"/> No		48. Ha Perdido la Gratuidad: <input type="checkbox"/> Si <input type="checkbox"/> No <input type="checkbox"/> NA
49. Ha realizado Prácticas Preprofesionales: <input type="checkbox"/> Si <input type="checkbox"/> No		50. Nro. Horas de Prácticas Preprofesionales Realizadas: <input type="text"/>



51. Tipo de Institución en el que realiza Prácticas Pre profesionales: <input type="checkbox"/> Pública <input type="checkbox"/> Privada <input type="checkbox"/> ONG <input type="checkbox"/> Otros <input type="checkbox"/> NA	
52. Sector Económico en el que realiza Prácticas Pre profesionales: <input type="text"/> (Escribir el número de acuerdo al detalle)	
1. Agricultura, ganadería, silvicultura y pesca 2. Explotación de minas y canteras 3. Industrias manufactureras 4. Suministro de electricidad, gas, vapor y aire acondicionado 5. Distribución de agua, alcantarillado, gestión de desechos y actividades de saneamiento. 6. Construcción 7. Comercio al por mayor y al por menor reparación de vehículos automotores y motocicletas 8. 9. Transporte y almacenamiento 10. Actividades de alojamiento y de servicio de comidas. 11. Información y comunicación bienes y servicios para uso propio	12. Actividades financieras y de seguros 13. Actividades inmobiliarias 14. Actividades profesionales, científicas y técnicas 15. Actividades de servicios administrativos y de apoyo 16. Administración pública y defensa; planes de seguridad social de afiliación obligat. 17. Enseñanza 18. Actividades de atención de la salud humana y de asistencia. 19. Artes, entretenimiento y recreación. 20. Otras actividades de servicio. 21. Actividades de los hogares como productores de bienes y servicios para uso propio 22. NA
53. Ha participado en algún Proyecto de Vinculación con la Sociedad en el Instituto: <input type="checkbox"/> Si <input type="checkbox"/> No	
54. Cuál es el alcance del Proyecto de Vinculación con la Sociedad: <input type="checkbox"/> Nacional <input type="checkbox"/> Provincial <input type="checkbox"/> Cantonal <input type="checkbox"/> Parroquial <input type="checkbox"/> NA	
55. El Estudiante se encuentra dedicado a: <input type="checkbox"/> Solamente al estudio <input type="checkbox"/> Trabaja y estudia	
56. Cuál es el nombre de la empresa donde labora:	
57. Indique cuál es el Sector Económico de la empresa: <input type="text"/> (Escribir el número de acuerdo al detalle)	
1. Agricultura, ganadería, silvicultura y pesca 2. Explotación de minas y canteras 3. Industrias manufactureras 4. Suministro de electricidad, gas, vapor y aire acondicionado 5. Distribución de agua, alcantarillado, gestión de desechos y actividades de saneamiento. 6. Construcción 7. Comercio al por mayor y al por menor reparación de vehículos automotores y motocicletas 8. 9. Transporte y almacenamiento 10. Actividades de alojamiento y de servicio de comidas. 11. Información y comunicación bienes y servicios para uso propio	12. Actividades financieras y de seguros 13. Actividades inmobiliarias 14. Actividades profesionales, científicas y técnicas 15. Actividades de servicios administrativos y de apoyo 16. Administración pública y defensa; planes de seguridad social de afiliación obligat. 17. Enseñanza 18. Actividades de atención de la salud humana y de asistencia. 19. Artes, entretenimiento y recreación. 20. Otras actividades de servicio. 21. Actividades de los hogares como productores de bienes y servicios para uso propio 22. NA
58. Para qué emplea sus ingresos económicos el estudiante: <input type="checkbox"/> Sustener sus estudios <input type="checkbox"/> Para mantener a su familia <input type="checkbox"/> Gastos personales <input type="checkbox"/> NA	
59. Usted o algún miembro de la familia recibe el bono de desarrollo humano: Si <input type="checkbox"/> No <input type="checkbox"/>	
60. Nivel de Formación del Padre: <input type="checkbox"/> Centro de Alfabetización <input type="checkbox"/> Jardín de Infantes <input type="checkbox"/> Primaria <input type="checkbox"/> Educación Básica <input type="checkbox"/> Secundaria <input type="checkbox"/> Educación Media <input type="checkbox"/> Superior no universitario <input type="checkbox"/> Superior Universitario <input type="checkbox"/> Post Grado <input type="checkbox"/> NA	
61. Nivel de Formación de la Madre: <input type="checkbox"/> Centro de Alfabetización <input type="checkbox"/> Jardín de Infantes <input type="checkbox"/> Primaria <input type="checkbox"/> Educación Básica <input type="checkbox"/> Secundaria <input type="checkbox"/> Educación Media <input type="checkbox"/> Superior no universitario <input type="checkbox"/> Superior Universitario <input type="checkbox"/> Post Grado <input type="checkbox"/> NA	
62. Ingresos del Hogar: \$	63. Número de Miembros del Hogar:

FIRMA ESTUDIANTE:

Anexo 3.- Muestra de base de datos utilizada en el estudio

Estudiante	Sexo	estadocivil	repitencia	ocupacion	ingresos	edad	carrera	retencion
1	Masculino	Soltero	0	0	0	26	PA	1
2	Femenino	Soltero	0	0	0	21	PA	1
3	Masculino	Soltero	0	1	0	23	PA	1
4	Femenino	Soltero	0	0	0	20	PA	1
5	Femenino	Soltero	0	0	1	25	PA	0
6	Masculino	Soltero	0	1	0	22	PA	1
7	Femenino	Soltero	0	0	0	22	PA	1
8	Femenino	Soltero	0	0	1	31	PA	1
9	Femenino	Soltero	1	1	0	24	PA	1
10	Femenino	Casado	0	0	1	22	PA	1
11	Femenino	Soltero	0	0	1	26	PA	1
12	Masculino	Soltero	1	0	1	26	PA	1
13	Femenino	Soltero	0	0	0	19	PA	1
14	Femenino	Soltero	0	0	0	23	PA	1
15	Femenino	Casado	0	1	0	23	PA	1
16	Femenino	Soltero	0	0	0	23	PA	1
17	Femenino	Soltero	0	0	0	20	PA	1
18	Femenino	Soltero	0	0	1	21	PA	0
19	Femenino	Divorciado	0	0	1	29	PA	1
20	Masculino	Soltero	0	0	0	21	PA	1
21	Masculino	Soltero	1	0	0	22	PA	1
22	Masculino	Soltero	1	1	1	22	PA	1
23	Masculino	Soltero	0	1	1	22	PA	0
24	Femenino	Soltero	0	0	1	22	PA	1
25	Femenino	Soltero	0	0	1	22	PA	1
26	Femenino	Soltero	0	1	0	27	PA	1
27	Femenino	Soltero	1	0	1	23	PA	1
28	Femenino	Soltero	0	0	1	24	PA	1
29	Femenino	Soltero	0	0	1	23	PA	1
30	Masculino	Soltero	0	0	1	22	PA	1
31	Femenino	Soltero	0	0	0	21	PA	1
32	Femenino	Soltero	1	0	0	22	PA	1
33	Femenino	Soltero	0	0	1	21	PA	0
34	Femenino	Soltero	0	0	0	21	PA	1
35	Masculino	Soltero	1	1	1	24	PA	0
36	Femenino	Soltero	0	0	0	20	PA	1
37	Femenino	Soltero	1	1	1	32	PA	1
38	Masculino	Soltero	0	0	0	20	PA	1
39	Masculino	Soltero	0	1	0	20	PA	1
40	Femenino	Casado	0	1	0	26	PA	1
41	Masculino	Soltero	0	0	1	23	PA	1

Estudiante	Sexo	estadocivil	repitencia	ocupacion	ingresos	edad	carrera	retencion
42	Femenino	Soltero	0	0	1	22	PA	1
43	Masculino	Soltero	0	0	1	22	PA	1
44	Femenino	Soltero	1	1	1	23	PA	1
45	Femenino	Soltero	0	0	1	25	PA	1
46	Femenino	Soltero	0	0	1	21	PA	1
47	Masculino	Soltero	0	0	0	27	PA	1
48	Femenino	Soltero	0	0	1	22	PA	0
49	Femenino	Soltero	0	0	1	23	PA	0
50	Femenino	Soltero	0	0	0	24	PA	1

Anexo 4.- CÓDIGO DE PROGRAMA R STUDIO DEL MODELO DE REGRESIÓN LOGÍSTICA

*****MODELO DE REGRESIÓN LOGÍSTICA*****

```
library(dplyr)
library(lmtest)
library(ResourceSelection)
library(MLmetrics)
library(InformationValue)
library(stargazer)
library(sjPlot)
library(DescTools)

mod1<-glm(formula = Deserción~Género+Estadocivil+Repitencia+
           Ocupación+Ingresos+Edad+Carrera,
           records, family=binomial(link = "logit"))

summary(mod1)
#Retirar variables no significativas de dos en dos, para ver como cambia modelo
#las variables a retirarse son las que tienen mayor valor p
#Se elimina género y edad
mod2<-glm(formula = Deserción~Estadocivil+Repitencia+
           Ocupación+Ingresos+Carrera,
           records, family=binomial(link = "logit"))
summary(mod2)
#Se elimina ingresos y edad
mod3<-glm(formula = Deserción~Repitencia+
           Ocupación+Carrera,
           records, family=binomial(link = "logit"))
summary(mod3)
#Se elimina ocupación
```

```
mod4<-glm(formula = Deserción~Repitencia+Carrera,  
           records, family=binomial(link = "logit"))  
summary(mod4)
```

```
anova(mod4, test = "Chisq")  
coef(mod4)  
exp(coef(mod4))
```

Anexo 5.- Código de programa R studio del **ÁRBOL DE DECISIÓN**

*****ÁRBOL DE DECISIÓN*****

```
library("rpart")
library("rattle")
library("rpart.plot")

#Modelo con Árbol de decisión

arbol<- rpart(data=data, formula = Deserción~Género+Estadocivil+Repitencia+
              Ocupación+Ingresos+Edad+Carrera,
              control = rpart.control(cp=0.01,minbucket = 30, minsplit = 20))
print(arbol)

#Gráfico del Árbol
fancyRpartPlot(arbol)
asRules(arbol)

#Prediciendo con arbol
pred_arbol<-predict(arbol)
arbol_pred<-cbind(data,pred_arbol)

ConfusionMatrix(pred_arbol,data$Deserción)
```

Anexo 6.- Certificado de entrega del trabajo de titulación al Instituto Superior
Tecnológico Luis A. Martínez Agronómico



La suscrita Rectora del Instituto Superior Tecnológico Luis A. Martínez (Agronómico) tiene a bien,

CERTIFICAR

Que: he recibido en conformidad el trabajo de investigación con el tema “Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica”, desarrollado por la Ing. Cristina Nataly Vinueza López, portadora de la CI N°1804486650, maestrante del Programa de Maestría en Matemática Aplicada-Cohorte Marzo 2019, que se desarrolló en la Facultad de Ingeniería en Sistemas, Electrónica e Industrial/Centro de Posgrados de la Universidad Técnica de Ambato.

Es todo cuanto puedo certificar en honor a la verdad, pudiendo la interesada hacer uso del presente documento en lo que estimare conveniente.

Ambato, 15 de enero del 2021.

MARYUXI
ALEJANDRA
CASTILLO
ORDNEZ
Lic. Maryuxi Castillo Mg.

Firmado digitalmente
por MARYUXI
ALEJANDRA CASTILLO
ORDNEZ
Fecha: 2021.01.16
21:04:29 -05'00'



Rectora del I.S.T Luis A. Martínez (Agronómico)

Av. Ruminahui y Pichincha Alta
Teléfono 032 419293
itslamagronomico@gmail.com