



**UNIVERSIDAD TÉCNICA DE AMBATO**

**FACULTAD DE INGENIERÍA EN SISTEMAS,  
ELECTRÓNICA E INDUSTRIAL**

**MAESTRÍA EN GESTIÓN DE BASES DE DATOS**

---

**Tema:** “LAS TÉCNICAS DE PREDICCIÓN Y SU INCIDENCIA EN LA DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL EN LA CARRERA DE DOCENCIA EN INFORMÁTICA DE LA FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACIÓN DE LA UNIVERSIDAD TÉCNICA DE AMBATO.”

---

Trabajo de Investigación, previo a la obtención del Grado Académico de  
Magíster en Gestión de Bases de Datos

**Autora:** Ingeniera Blanca Rocio Cuji Chacha

**Director:** Ingeniero Edison Homero Álvarez Mayorga, Magíster

**Ambato- Ecuador  
2016**

**A la Unidad Académica de Titulación de la Facultad de Ingeniería en Sistemas,  
Electrónica e Industrial**

El Tribunal receptor del Trabajo de Investigación presidido por Ingeniero José Vicente Morales Lozada, Magíster. e integrado por los señores Ingeniero Clay Fernando Aldás Flores, Magíster, Ingeniero Fabián Rodrigo Salazar Escobar, Magíster, Economista Andrés Fernando Cisneros Cadena, Magíster, designados por la Unidad Académica de Titulación de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial de la Universidad Técnica de Ambato, para receptor el Trabajo de Investigación con el tema: “LAS TÉCNICAS DE PREDICCIÓN Y SU INCIDENCIA EN LA DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL EN LA CARRERA DE DOCENCIA EN INFORMÁTICA DE LA FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACIÓN DE LA UNIVERSIDAD TÉCNICA DE AMBATO”, elaborado y presentado por la señora Ingeniera, Blanca Rocio Cuji Chacha, para optar por el Grado Académico de Magister Gestión de Bases de Datos; una vez escuchada la defensa oral del Trabajo de Investigación el Tribunal aprueba y remite el trabajo para uso y custodia en las bibliotecas de la UTA.

-----  
Ing. José Vicente Morales Lozada Mg.  
**Presidente del Tribunal**

-----  
Ing. Clay Fernando Aldás Flores Mg.  
**Miembro del Tribunal**

-----  
Ing. Fabián Rodrigo Salazar Escobar Mg.  
**Miembro del Tribunal**

-----  
Econ. Andrés Fernando Cisneros Cadena Mg.  
**Miembro del Tribunal**

## **AUTORÍA DEL TRABAJO DE INVESTIGACIÓN**

La responsabilidad de las opiniones, comentarios y críticas emitidas en el trabajo de titulación con el tema: “LAS TÉCNICAS DE PREDICCIÓN Y SU INCIDENCIA EN LA DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL EN LA CARRERA DE DOCENCIA EN INFORMÁTICA DE LA FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACIÓN DE LA UNIVERSIDAD TÉCNICA DE AMBATO”, le corresponde exclusivamente a: Ingeniera Blanca Rocio Cuji Chacha, Autora bajo la Dirección del Ingeniero Edison Homero Álvarez Mayorga, Magister, Director del trabajo de Investigación; y el patrimonio intelectual a la Universidad Técnica de Ambato.

.....

Ing. Blanca Rocio Cuji Chacha

C.c. 1803127594

**AUTORA**

.....

Ing. Edison Homero Álvarez Mayorga, Mg.

c.c. 1801225960

**DIRECTOR**

## **DERECHOS DE AUTOR**

Autorizo a la Universidad Técnica de Ambato, para que el Trabajo de Investigación, 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, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad.

.....  
Ing. Blanca Rocio Cuji Chacha

C.C. 1803127594

## ÍNDICE GENERAL DE CONTENIDOS

Portada.....	i
A la Unidad Académica de Titulación de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial .....	ii
AUTORÍA DEL TRABAJO DE INVESTIGACIÓN.....	iii
DERECHOS DE AUTOR.....	iv
DEDICATORIA .....	xiii
AGRADECIMIENTO.....	xiv
ÍNDICE GENERAL DE CONTENIDOS.....	v
ÍNDICE DE TABLAS .....	ix
ÍNDICE DE GRÁFICOS .....	xii
RESUMEN EJECUTIVO .....	xv
EXECUTE SUMMARY .....	xvi
INTRODUCCIÓN .....	1
<b>CAPÍTULO I.....</b>	<b>2</b>
<b>EL PROBLEMA .....</b>	<b>2</b>
1.1. Tema.....	2
1.2. Planteamiento del Problema.....	2
1.1.1. Contextualización.....	2
1.1.2. Análisis Crítico.....	6
1.1.3. Árbol de Problemas.....	7
1.1.4. Prognosis .....	7
1.1.5. Formulación del Problema .....	7
1.1.6. Interrogantes.....	8
1.1.7. Delimitación de la Investigación.....	8

1.1.7.1.	Delimitación de Contenidos.....	8
1.1.7.2.	Delimitación Temporal.....	8
1.1.7.3.	Delimitación Espacial.....	8
1.1.7.4.	Unidades de Observación.....	8
1.3.	Justificación.....	9
1.4.	Objetivos.....	9
1.1.8.	Objetivo General.....	9
1.1.9.	Objetivos Específicos.....	9
<b>CAPÍTULO II.....</b>		<b>11</b>
<b>MARCO TEÓRICO.....</b>		<b>11</b>
2.1.	Antecedentes Investigativos.....	11
2.2.	Fundamentación Filosófica.....	13
2.3.	Fundamentación Legal.....	13
2.4.	Categorías Fundamentales.....	14
2.5.	Hipótesis.....	23
2.6.	Señalamiento de Variables.....	23
2.6.1.	Variable Independiente.....	23
2.6.2.	Variable Dependiente.....	23
<b>CAPÍTULO III.....</b>		<b>24</b>
<b>METODOLOGÍA.....</b>		<b>24</b>
3.1.	Enfoque de la Investigación.....	24
3.2.	Modalidad Básica de la Investigación.....	24
3.3.	Niveles o Tipos de Investigación.....	25
3.3.1.	Exploratorio.....	25

3.3.2. Descriptivo .....	25
3.3.3. Correlacional .....	25
3.4. Población y muestra .....	25
3.5. Operacionalización de Variables.....	27
3.6. Técnicas e instrumentos de recolección de información.....	31
3.7. Procesamiento de la Información.....	31
3.8. Validez y Confiabilidad .....	31
3.8.1. Validez .....	31
3.8.2. Confiabilidad.....	32
3.9. Plan de recolección de la información .....	32
3.10. Plan de procesamiento de la información .....	32
<b>CAPÍTULO IV .....</b>	<b>34</b>
<b>ANÁLISIS E INTERPRETACIÓN DE RESULTADOS.....</b>	<b>34</b>
4.1. Análisis de Resultados .....	34
4.2. Verificación de la Hipótesis .....	84
<b>CAPÍTULO V.....</b>	<b>90</b>
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>90</b>
5.1. Conclusiones .....	90
5.2. Recomendaciones.....	91
<b>CAPÍTULO VI .....</b>	<b>92</b>
<b>PROPUESTA.....</b>	<b>92</b>
6.1. Tema:.....	92
6.2. Datos Informativos.....	92

6.3. Antecedentes de la propuesta .....	92
6.4. Justificación.....	93
6.5. Objetivos de la Propuesta.....	94
Objetivos Específicos.....	94
6.6. Análisis de Factibilidad.....	94
6.7. Fundamentación .....	95
6.8. Metodología .....	97
6.8.1. Conclusiones de la propuesta .....	118
6.8.2. Recomendaciones de la propuesta .....	119
<b>BIBLIOGRAFÍA.....</b>	<b>124</b>
<b>ANEXOS .....</b>	<b>127</b>
Anexo 1: Formato de encuesta.....	127



## ÍNDICE DE TABLAS

Tabla 1: Población.....	26
Tabla 2: Operacionalización Variable Independiente .....	27
Tabla 3: Operacionalización Variable Dependiente .....	29
Tabla 4: Edad Estudiantes.....	34
Tabla 5: Lugar de nacimiento-Estudiantes.....	35
Tabla 6: Genero- Estudiantes .....	35
Tabla 7: Estado Civil.....	36
Tabla 8: Discapacidad .....	36
Tabla 9: Identificación Étnica .....	37
Tabla 10: Razones que propician la Deserción Estudiantil.....	37
Tabla 11: Automatización de procesos .....	39
Tabla 12: Técnicas de Predicción .....	39
Tabla 13: Culminar Semestre.....	40
Tabla 14: Dedicación trabajo autónomo .....	41
Tabla 15: Decisión de tomar una carrera universitaria .....	42
Tabla 16: Enfermedad .....	42
Tabla 17: Nivel Económico .....	43
Tabla 18: Gastos de estudio .....	44
Tabla 19: Ayuda para realizar estudios .....	45
Tabla 20: Porcentaje de deserción.....	45
Tabla 21: Nivel de abandono de estudios .....	46
Tabla 22: Nivel de estudios de los padres.....	47
Tabla 23: Nivel de estudios del padre .....	46

Tabla 24: Espacios para estudiar.....	48
Tabla 25: Estimulo para la realización de estudios.....	498
Tabla 26: Conocimientos útiles.....	49
Tabla 27: Dificultad den las asignaturas .....	50
Tabla 28: Examen de admisión.....	51
Tabla 29: Reprobación de asignaturas .....	52
Tabla 30: Abandono de estudios .....	53
Tabla 31: Edad - Estudiantes desertores .....	54
Tabla 32: Lugar de nacimiento- Estudiantes desertores .....	54
Tabla 33: Género.....	55
Tabla 34: Estado Civil- Estudiantes desertores.....	55
Tabla 35: Capacidades especiales- Estudiantes desertores .....	56
Tabla 36: Etnia .....	56
Tabla 37: Razones para la Deserción Estudiantil -Estudiantes Desertores.....	57
Tabla 38: Automatización de procesos-Estudiantes Desertores .....	58
Tabla 39: Técnicas de Minería de datos-Estudiantes Desertores.....	59
Tabla 40: Año de deserción-Estudiantes Desertores.....	59
Tabla 41: Trabajo Autónomo-Estudiantes Desertores .....	59
Tabla 42: Carrera Universitaria -Estudiantes Desertores.....	61
Tabla 43: Enfermedad-Estudiantes Desertore.....	62
Tabla 44: Nivel económico de la familia-Estudiantes Desertores .....	62
Tabla 45: Gastos Estudiantiles-Estudiantes Desertores .....	63
Tabla 46: Becas -Estudiantes Desertores .....	64
Tabla 47: Porcentaje de deserción -Estudiantes Desertores.....	65
Tabla 48: Créditos tomado -Estudiantes Desertores .....	66
Tabla 49: Estudios de los padres-Madre .....	67

Tabla 50: Estudio de los padres-Padre .....	67
Tabla 51: Espacio adecuado-Estudiantes Desertores .....	68
Tabla 52: Fomento de estudios-Estudiantes Desertores.....	69
Tabla 53: Conocimientos adquiridos-Estudiantes Desertores .....	69
Tabla 54: Dificultad en las asignaturas-Estudiantes Desertores .....	70
Tabla 55: Examen de admisión .....	71
Tabla 56: Materias reprobadas-Estudiantes Desertores .....	72
Tabla 57: Abandono de estudios-Estudiantes Desertores .....	72
Tabla 58: Edad-Docentes .....	73
Tabla 59: Razones de Deserción Estudiantil-Docentes .....	74
Tabla 60: Automatización de procesos -Docentes.....	75
Tabla 61: Técnicas de minería de datos-Docentes.....	75
Tabla 62: Culminación semestre-Docentes.....	76
Tabla 63: Trabajo Autónomo-Docentes.....	77
Tabla 64: Nivel Económico-Docentes .....	78
Tabla 65: Recursos Económicos .....	79
Tabla 66: Becas -Docentes.....	80
Tabla 67: Porcentaje de deserción-Docentes .....	80
Tabla 68: Créditos tomados .....	81
Tabla 69: Conocimientos adquiridos -Docentes .....	82
Tabla 70: Estudiantes Desertores-Docentes.....	83

## ÍNDICE DE GRÁFICOS

Figura 1: Datos personales .....	100
Figura 2 : Notas .....	100
Figura 3: Integración de datos .....	101

## **DEDICATORIA**

A mis padres Francisco y Transito

A mis estudiantes de la Carrera de Docencia en Informática

## **AGRADECIMIENTO**

A la Universidad Técnica de Ambato, a la Facultad de Ciencias Humanas y de la Educación y por su intermedio a la Carrera de Docencia en Informática por prestar las facilidades necesarias para la elaboración del proyecto; de igual manera al Ing. Edison Álvarez, tutor del proyecto de investigación por su colaboración.

**UNIVERSIDAD TÉCNICA DE AMBATO**  
**FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA**  
**E INDUSTRIAL**  
**MAESTRÍA EN GESTIÓN DE BASES DE DATOS**

**TEMA:**

“LAS TÉCNICAS DE PREDICCIÓN Y SU INCIDENCIA EN LA DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL EN LA CARRERA DE DOCENCIA EN INFORMÁTICA DE LA FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACIÓN DE LA UNIVERSIDAD TÉCNICA DE AMBATO”

**Autora:** Ing. Blanca Rocío Cuji Chacha

**Director:** Ing. Edison Homero Alvarez Mayorga, Mg.

**Fecha:** 21 de julio 2016

**RESUMEN EJECUTIVO**

El estudio sobre Deserción Estudiantil tuvo como objetivo primario detectar el porcentaje de abandono escolar que presenta la Carrera de Docencia en Informática, a partir del año 2006 hasta el año 2015, con base en ésta información se aplicó el algoritmo de Árboles de decisión para diseñar un prototipo de modelo predictivo de Deserción Estudiantil, la metodología usada se basa en el método KDD(*Knowledge Discovery in Database*), detallado en cinco etapas , selección, procesamiento, transformación, minería de datos, e interpretación de la información.

Posterior a la aplicación del algoritmo se obtuvo un árbol de decisión de cuatro niveles de profundidad, evidenciando que las variables nivel y notas tienen mayor influencia en la Deserción Estudiantil dentro de la Carrera.

Finalmente se obtuvieron cuatro reglas que fueron programadas y visualizadas en una interfaz web, que evalúa a los nuevos posibles desertores de la Carrera de Docencia en Informática.

**Descriptor:** Árboles de decisión, Algoritmo de Clasificación, Base de datos, Deserción Estudiantil, Minería de Datos, Modelo, Método KDD, Predicción, Toma de decisiones, Visualización de datos.

**UNIVERSIDAD TÉCNICA DE AMBATO**  
**FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA**  
**E INDUSTRIAL**  
**MAESTRÍA EN GESTIÓN DE BASES DE DATOS**

**Theme:**

“Prediction Techniques and their influence in the detection of Student Dropout rate in the Computer Science Teaching Program of the Humans Sciences and Education Faculty of the Technical University of Ambato”

**Author:** Ing. Blanca Rocío Cuji Chacha

**Directed by:** Ing. Edison Homero Alvarez Mayorga, Mg.

**Date:** July 26, 2016

**EXECUTE SUMMARY**

The present study about Student Dropout had as the main goal to determine the percentage of student dropout from the Computer Science Teaching program. The time frame for the study started in 2006 and finished in 2015. In order to process the information gathered we used the Decision Trees algorithm to create a prototype model to predict the percentage of student dropouts. The methodology is based on the Knowledge Discovery in Database(KDD)which has five stages: selection, processing, transformation, data mining, and interpretation. Afterwards, we got a Decision tree with four levels. The variables level and grades are very influent in the student dropout average.

Finally, we obtained four rules that were programmed and visualized in a web interface, which assesses the future possible droppers.

**Keywords:** Classification Algorithm, Data visualization, Database, Data Mining, Decision Making, Decision trees, Student Dropout, KDD Method, Model, Prediction.



## INTRODUCCIÓN

La investigación sobre: “Las Técnicas de Predicción y su incidencia en la Detección de Patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato”. Se fundamenta en la necesidad que tiene la Facultad de bajar los índices de deserción escolar y por ende mejorar la retención estudiantil.

Está estructurado por los siguientes capítulos:

**Capítulo I:** Denominado EL PROBLEMA se analiza el problema de investigación en diferentes contextos macro, meso y micro, se formularon interrogantes que guiaron el proceso investigativo, además se delimito el objeto de investigación, la justificación y elaboración de objetivos generales y específicos.

**Capítulo II:** Denominado MARCO TEÓRICO contiene los antecedentes investigativos, fundamentaciones filosófica, tecnológica y legal, se plantea una hipótesis de trabajo general y se señalan sus variables.

**Capítulo III:** Denominado METODOLOGÍA se propone un enfoque, modalidad y nivel de investigación, se determina la población y el tamaño de la muestra, se operacionaliza las variables, y se plantea la manera en que se llevará a cabo la recolección de la información, así como su procesamiento y análisis de la misma.

**Capítulo IV:** Denominado ANÁLISIS E INTERPRETACIÓN DE RESULTADOS, contiene la representación gráfica y en tablas de la tabulación, análisis e interpretación de los datos tomados a los estudiantes regulares, desertores y docentes de la Carrera de Docencia en Informática, así como la comprobación de la hipótesis planteada utilizando técnicas estadísticas.

**Capítulo V:** Denominado CONCLUSIONES Y RECOMENDACIONES, muestra las conclusiones y recomendaciones que llevo la investigación.

**Capítulo VI:** Denominado PROPUESTA presenta el modelo de predicción de Deserción Estudiantil para la Carrera de Docencia en Informática.

# **CAPÍTULO I**

## **EL PROBLEMA**

### **1.1. Tema**

Las Técnicas de Predicción y su incidencia en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato.

### **1.2. Planteamiento del Problema**

#### **1.2.1 Contextualización**

La conceptualización de deserción escolar difiere entre varios autores, sin embargo, si una institución educativa pierde a sus alumnos se evidenciará en sus índices de retención estudiantil. El abandono o deserción de un estudiante en un determinado nivel puede basarse en la no matriculación de dichos estudiantes al siguiente nivel (Estellés, Alcover, Dapena-Janeiro, Valderruten, Satorre, & LLopis, 2009).

Por otro lado, una definición más común es la propuesta por (González González, 2006) quien sostiene que la deserción escolar es la ausencia definitiva y sin causa justificada de la institución educativa por parte del alumno sin haber finalizado la etapa educativa que se esté cursando.

Para muchos alumnos abandonar su proceso educativo detiene su vida académica y reduce su participación en un campo laboral altamente competitivo.

Según (Zhang, 2000) los índices de reprobación y deserción en la mayoría de las instituciones educativas del nivel medio y superior del mundo muestran valores elevados con relación a los resultados obtenidos en la educación básica, donde la gran mayoría de los estudiantes aprueban. Algunos de los indicadores de mayor importancia, que revela la calidad que tiene una institución educativa es el porcentaje de reprobación y deserción escolar junto con el número de estudiantes matriculados.

Por otro lado, las técnicas de predicción permiten extraer conocimiento no implícito de los datos, la Minería de Datos (Data Mining, DM) es un campo multidisciplinar que resulta beneficioso al momento de descubrir patrones o diseñar modelos de predicción.

Dentro de la minería de datos se pueden considerar dos enfoques como técnicas de predicción: las clásicas técnicas estadísticas como la regresión lineal, el análisis factorial y por otro lado la de **minería de datos** éstas proceden en su gran mayoría de las contribuciones hechas por la Inteligencia artificial como las redes neuronales, lógica difusa, y mecanismos de aprendizaje autónomo que utilizan algoritmos como Clustering, clasificación, arboles de decisión entre otros.(Aldehuela, 2015), las dos técnicas se aplican en la extracción, predicción, clasificación y análisis de información para tener una mayor precisión en la toma de decisiones.

Así pues, es menester analizar varios estudios realizados sobre las técnicas de predicción y Deserción Estudiantil en diferentes países de América y el mundo.

En México, autores como (Martínez & Pérez, 2008) en su trabajo sobre “efecto de la trayectoria Académica en el Desempeño de Estudiantes de Ingeniería”, se realiza un análisis en términos de permanencia y de eficiencia terminal, tomando como población a los estudiantes de las cohortes que ingresaron al programa de Ingeniería Civil entre los años 1990 y 1996, en donde se toma, los resultados del examen de admisión y la eficiencia de egreso de la cohorte para predecir quienes concluirán el programa y quienes abandonarán prematuramente sus estudios a lo

cual se propone un modelo de predicción a partir del análisis estadístico del historial académico de los alumnos.

En España se encuentran trabajos relacionados con el tema como: “Predicción del fracaso y el abandono escolar mediante técnicas de Minería de Datos”(Márquez, 2015), en donde se aborda el problema del fracaso escolar de los estudiantes se propone una metodología para predecir a los alumnos que se encuentran en riesgo de abandonar o reprobado sus estudios utilizando diferentes técnicas de Minería de Datos.

Es necesario destacar trabajos como “Predicting School Failure using Data Mining” de México donde se trabaja con 670 estudiantes para predecir la deserción estudiantil basado en varios experimentos que permiten mejorar la precisión en la predicción del rendimiento final de los estudiantes (Marquez & Romero, 2011).

Por otro lado en el trabajo “Student Dropout Analysis with Application of Data Mining Methods” en donde se analiza a 715 estudiantes que se agrupan en categorías en función de su rango en la inscripción. Con base en el análisis realizado en SPSS, se grafican los datos que son seleccionados para describir la estructura de los estudiantes (Jadrić, Garača, & Željko, 2010).

La Constitución Política del Ecuador en su sección quinta Art. 28 manifiesta: “La educación pública será universal y laica en todos sus niveles, y **gratuita** hasta el tercer nivel de educación superior...”.

La inversión en la gratuidad de la educación puede representar un perjuicio para el estado. Carlos Cedeño, rector de la Universidad de Guayaquil manifiesta: "Apenas el 30% de los bachilleres que ingresa se gradúa. Quiere decir que un 70% se va quedando en el camino, y eso, es grave porque el Estado está invirtiendo en ellos".

Ese problema también se resalta en estudios hechos por la Unesco (entre 2000 y 2005) que señala que América Latina pierde hasta \$420 millones por la deserción. Dicho en declaraciones al Diario hoy, 2008 fecha del periódico.

Según Miguel Tapia (2005), presidente del Movimiento Integración de la Universidad Central, en esta institución existen aproximadamente 11 mil estudiantes asegura que se inscriben al año 1500 aspirantes, pero solo 200 se gradúan.

Por otro lado, Mónica Chávez (2005) catedrática de la Escuela Superior Politécnica del Litoral (ESPOL), se determinó que el 10,59% de los estudiantes de ingeniería desertaron de sus estudios universitarios siendo en su gran mayoría estudiantes de género masculino de edades comprendidas entre 19 y 22 años.

María José Rubio docente de la Pontificia Universidad Católica del Ecuador sede Ibarra en su estudio sobre la inserción y deserción laboral manifiesta que aproximadamente un 13,4% en promedio de los estudiantes de las facultades de esta universidad desertan de sus estudios (Rubio, Tocaín, & Mantila, 2012).

Estadísticas publicadas por Diario Hoy (2008) muestran la postura de los estudiantes a la carrera que eligió y al deseo de retirarse de la especialidad que eligieron



**Gráfico: Encuesta de Deserción Estudiantil**  
Fuente: Diario Hoy, noviembre 2008

Universidad Técnica de Ambato como las diferentes universidades del país pasa por el mismo fenómeno de Deserción Estudiantil, muy pocos datos existen sobre la totalidad de desertores de las diferentes facultades, es poco evidenciado las causas que generan este fenómeno, existen escasos medios automatizados para determinar las variables que generan la deserción, y los indicadores que conllevarían atacar este problema.

### 1.2.2 Análisis Crítico

La Ley Orgánica de Educación Superior (LOES) en la Disposición Transitoria constitucional vigésima dispone que *“Todas las instituciones de educación superior, así como sus carreras, programas y posgrados deberán ser evaluados y acreditados conforme a la ley.”*, basado en este criterio y con motivos de evaluación a nivel de carreras, a partir del año 2012 se llevaron a cabo varios estudios manuales en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación, a fin de poder medir la tasa de deserción y retención estudiantil, este proceso se realiza en forma manual, es tedioso y requiere de tiempo.

La información que se encuentra en la secretaría de la carrera no está depurada, y debido al número de cohortes que se requiere analizar el proceso manual es largo y repetitivo, los informes presentados solo muestran estadísticas eventuales sobre la deserción, no se determinan variables ni se establecen patrones que permitan la futura toma de decisiones.

Uno de los objetivos Estratégicos Institucionales de la UTA es:

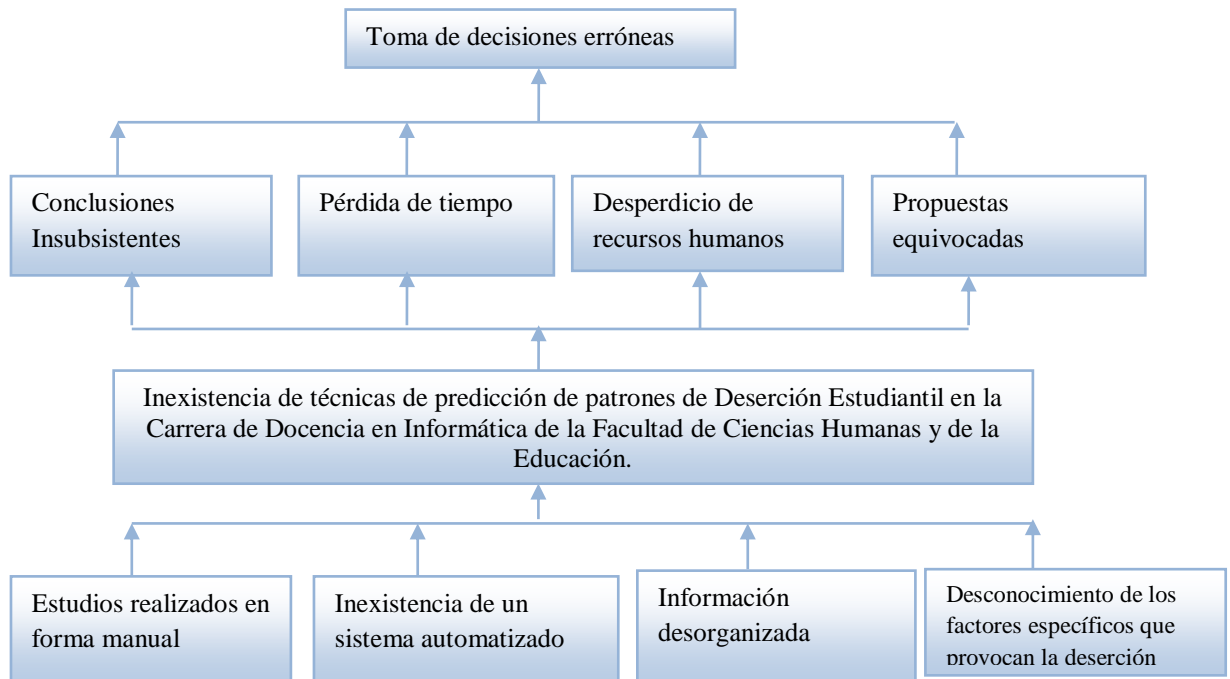
*“Formar y especializar profesionales con liderazgo, responsabilidad social ambiental con sólidos conocimientos científicos, tecnológicos y artísticos, que entiendan la realidad socioeconómica del Ecuador”.*

Según los estudios realizados manualmente por parte de la comisión interna de evaluación y acreditación, y del proceso de seguimiento a graduados de la Carrera de Docencia en Informática, la eficiencia terminal es del 18,73% para el año 2013, sin encontrarse evidencia sobre estudios realizados acerca de la tasa de Deserción Estudiantil, en los dos últimos semestres a partir del año 2015.

Como se evidencia en las estadísticas presentadas el objetivo estratégico institucional ve mermado su cumplimiento debido al bajo índice de graduados en la Carrera, la Facultad por su parte no cuenta con un sistema automatizado que

permita detectar a tiempo las causas que con mayor frecuencia provocan una deserción, lo que impide prever un posible abandono escolar.

### 1.2.3 Árbol de Problemas



### 1.2.4 Prognosis

Es imperiosa la necesidad de mostrar datos numéricos que revelen las causas de una Deserción Estudiantil, así como patrones y modelos que determinen las variables que intervienen en éste fenómeno, no solo con motivos de evaluación sino principalmente para atacar en un futuro las causas y disminuir los índices de abandono escolar.

### 1.2.5 Formulación del Problema

¿De qué manera las Técnicas de Predicción inciden en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato?

## **1.2.6 Interrogantes**

- ¿Cuál es la situación actual de la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación, con respecto a la Deserción Estudiantil?
- ¿Qué técnicas de predicción se aplican actualmente en la Facultad de Ciencias Humanas para detectar una posible Deserción Estudiantil?
- ¿Es necesario establecer las características del modelo que permitirá detectar patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación?

## **1.2.7 Delimitación de la Investigación**

### **1.2.7.1 Delimitación de Contenidos**

**Contenido:** Informática

**Área:** Base de Datos

**Aspecto:** Técnicas de Predicción - Minería de Datos

### **1.2.7.2 Delimitación Temporal**

La investigación se llevó a cabo en el año 2015

### **1.2.7.3 Delimitación Espacial**

Con la presente investigación se diagnosticó, analizó, selecciono y utilizo la minería de datos como técnica de predicción para detectar patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato.

### **1.2.7.4 Unidades de Observación**

Estudiantes de la Carrera de Docencia en Informática

Estudiantes desertores de la Carrera de Docencia en Informática



### **1.3. Justificación**

La aplicación de una técnica de predicción como la minería de datos permitió visualizar las causas e identificar patrones por las cuales los estudiantes abandonan sus estudios superiores, expresados mediante estadísticas, los porcentajes de deserción por curso y carrera, utilizando cohortes para su estudio, esto fue posible mediante la automatización de procesos. Mediante un análisis de las gráficas estadísticas que muestren los diferentes factores económicos, sociales, etc., las autoridades de la institución podrán tomar decisiones para prever posibles abandonos estudiantiles.

Es factible pues la investigación se financió con fondos propios del autor y el apoyo de las autoridades de la Facultad, se cuenta con la bibliografía documental apropiada que está directamente relacionada con el tema de investigación, la misma que servirá de apoyo para extraer los fundamentos teóricos y relacionarlos con la práctica.

Los Beneficiarios directos son 485 estudiantes de la Carrera de Docencia en Informática desde el año 2006, y de forma indirecta las autoridades de la institución, los docentes, el personal administrativo como secretarias.

### **1.4. Objetivos**

#### **1.4.1 Objetivo General**

Determinar cómo las Técnicas de Predicción inciden en la Detección de Patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato

#### **1.4.2 Objetivos Específicos**

- Estudiar la situación actual sobre la Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación.

- Analizar las Técnicas de Predicción que se aplican actualmente en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación, para detectar la Deserción Estudiantil.
- Establecer las características del modelo que permitirá detectar patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación.

## CAPÍTULO II

### MARCO TEÓRICO

#### 2.1. Antecedentes Investigativos

En el estudio sobre “*Análisis de Deserción –Permanencia de estudiantes Universitarios utilizando técnicas de clasificación en Minería de Datos*” de (Eckert & Suénaga, 2015) se analiza información sobre la Deserción Estudiantil, basado en datos personales y antecedentes educativos de los estudiantes de la Carrera de Ingeniera en Informática de la Universidad Gastón Dachary en Argentina, el autor utiliza para la selección y depuración de datos aplicación de algoritmos de clasificación como árboles de decisión y redes bayesianas. El estudio presenta como variables de influencia en la deserción, la edad de ingreso, el establecimiento educativo del cual proviene el estudiante, lugar de origen (localización geográfica), número de desertores por año, cantidad de materias cursadas por año, número de aprobados por año, con base en estos parámetros el autor llega a las siguientes conclusiones.

“El proceso de preparación y acondicionamiento de los datos es la etapa más extensa y a la vez fundamental porque en gran medida los resultados posteriores dependen de ésta. En las etapas intermedias, es crítico llevar a cabo análisis e interpretaciones de resultados parciales, ya que a partir de éstos se retoma el proceso y continúa la depuración y refinamiento del conocimiento extraído.

... durante el primer año de la carrera es donde adquieren mayor importancia las acciones de contención, apoyo, tutoría y todas aquellas actividades que mejoren la situación académica del alumno al ingreso en la universidad.

Se detectaron atributos que al procesarlos y asociarlos a criterios específicos, se relacionan fuertemente con la deserción y permanencia, el principal de ellos es cantidad de asignaturas aprobadas del primer año, debido a que marca una tendencia notable sobre el resto de la carrera; otros que se destacan son: el número de asignaturas cursadas, los casos donde el estudiante no regulariza la materia al cursarlas, la edad de ingreso, la procedencia; la combinación de estos criterios obtuvo porcentajes de aciertos, de entre un 76% y un 80% de los casos clasificados correctamente.”(Eckert & Suénaga, 2015).

Por otro lado (Salazar, Gosálvez, Miralles, & Vergara, 2016) en su trabajo sobre “*A case study of knowledge discovery on academic achievement, student desertion and student retention*” presenta un estudio sobre el descubrimiento de conocimiento basado en el análisis de datos académicos, basado en información sobre el rendimiento, la deserción y retención de los estudiantes usando clustering y reglas de decisión de datos. Se utilizaron técnicas de minería como el algoritmo C-mean con estadísticas homogéneas aplicados a grupos descritos cualitativamente, legando a las siguientes conclusiones:

- El estudio cubre seis etapas incluyendo pre-procesamiento de datos, análisis estadístico, agrupación de procesos, presentación del conocimiento a través de la representación de los árboles de decisión, la interpretación y evaluación de los, así como una acción estratégica preliminar.
- Es necesario añadir más variables de estudio, por ejemplo, el horario del curso, metodología y número de estudiantes; currículo profesor y curso evaluación; y la información adicional acerca de la estudiante como el estado civil, las actividades de investigación.

De igual manera (Gerben, Mykola, & Jan, 2009) en su investigación sobre “*Predicting Students Drop Out: A Case Study*” manifiesta la necesidad de la supervisión y el apoyo que se debe dar a los estudiantes en el primer año de universidad, además describe los resultados sobre el modelo de predicción de

abandono escolar realizado con los estudiantes de ingeniería eléctrica, identifica los factores que influyen, se muestran resultados experimentales utilizando técnicas de minería de datos como clasificadores (árboles de decisión ) con una precisión entre 75 y 80 %. En el estudio se llega a las siguientes conclusiones:

- Se ha demostrado que el aprendizaje en cuanto a costo-sensible ayuda a la determinación de errores de clasificación
- La asignatura de Álgebra Lineal, es decisiva en la decisión de los estudiantes.

## **2.2. Fundamentación Filosófica**

Básicamente la presente investigación se sustenta en una teoría constructivista, generadora de nuevo conocimiento basado en datos históricos, opiniones pasadas y situaciones actuales, mediante un análisis de la información.

## **2.3. Fundamentación Legal**

Esta investigación se fundamenta en las siguientes leyes y reglamentos:

### **PLAN NACIONAL BUEN VIVIR**

#### **OBJETIVO 5: Planificando el Futuro**

5.1.1. Uno de los grandes retos del Buen Vivir es mejorar la calidad de vida de los ecuatorianos.

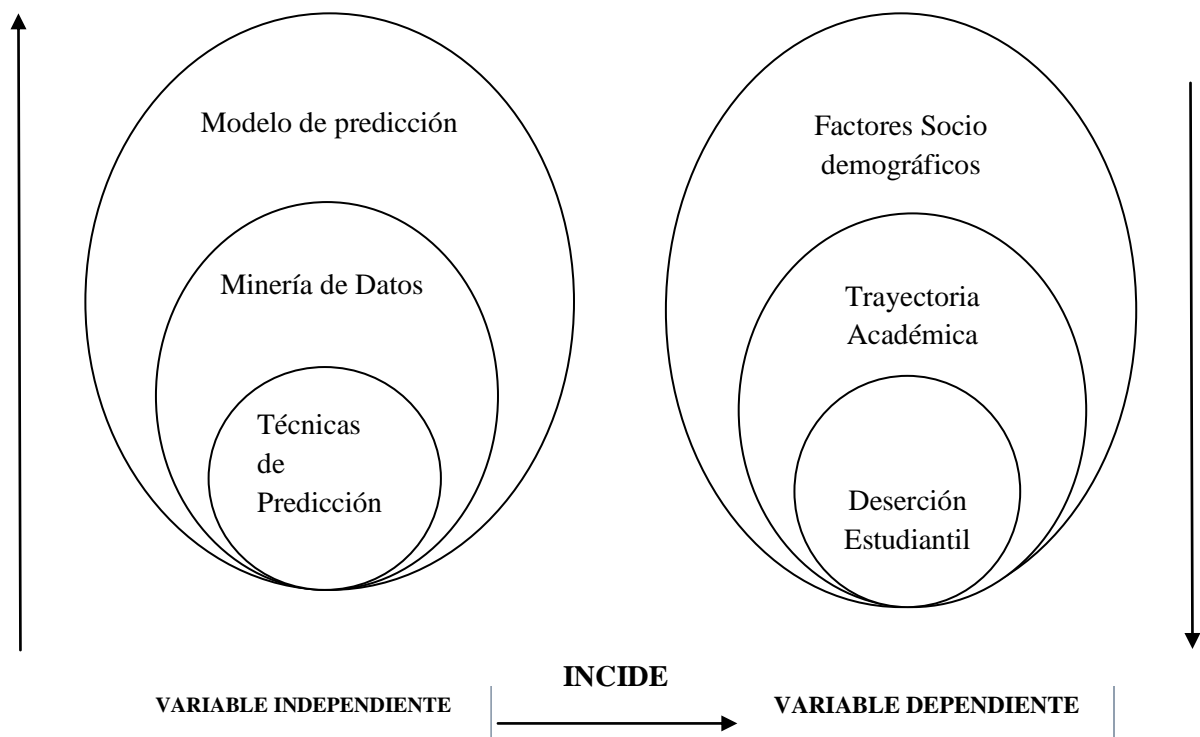
Para lograrlo, la diversificación productiva y el crecimiento de la economía deben dirigirse al cumplimiento progresivo de los derechos en **educación**, salud, empleo y vivienda, la reducción de la inequidad social, y la ampliación de las capacidades humanas en un entorno participativo y de creciente cohesión social, con respeto a la diversidad cultural.

## OBJETIVO 9: Garantizar el Trabajo digno en todas sus formas

**Literal e.:** Priorizar el desarrollo de iniciativas económicas vinculadas al desarrollo de las TICs, aprovechando las capacidades desarrolladas en software y de los recursos de la biodiversidad, creando espacios e infraestructura pertinente, que sustenten su productividad.

**Literal g:** Apoyar el desarrollo de los emprendimientos juveniles en ámbitos de financiamiento, capacitación, transferencia tecnológica y gestión.

### 2.4. Categorías Fundamentales



### Técnicas de predicción

Existen dos enfoques o áreas dentro del Data Mining: las técnicas clásicas de Estadística y las de minería de datos propiamente dicha que derivan, en su mayoría, de las aportaciones de la Inteligencia Artificial como las redes

neuronales, lógica borrosa, mecanismos de aprendizaje automático (Aldehuela, 2015).

Las funciones estadísticas ofrecen diversos métodos de pronóstico para dar apoyo a la toma de decisiones empresariales. Aunque no son propias de minería de datos (se utilizaban mucho antes de que surgiera este concepto), resultan de gran utilidad a la hora de descubrir patrones o elaborar modelos de predicción. Pueden emplearse para obtener más información sobre los datos, lo que permitirá tomar decisiones más acertadas cuando se apliquen los procesos de minería. Algunos ejemplos de estas técnicas son la regresión lineal, el Análisis Factorial, el Análisis de Componentes Principales. Mientras que entre las técnicas de minería destacan el Clustering, la Clasificación, la predicción. (Aldehuela, 2015).

### **Minería de Datos**

La minería de datos es entendida como el proceso de descubrir conocimientos interesantes, como patrones, asociaciones, cambios, anomalías y estructuras significativas a partir de grandes cantidades de datos almacenadas en bases de datos, datawarehouses, o cualquier otro medio de almacenamiento de información. (BRITOS & HOSSIAN, 2005).

La aplicación de algoritmos de minería de datos requiere de actividades previas destinadas a preparar los datos de manera homogénea. Esta primera etapa es también conocida como ETL (Extract, Transform and Load) (KIMBALL, 2002). Un proceso completo de aplicación de técnicas de minería, mejor conocido como proceso de descubrimiento del conocimiento en bases de datos establece a la minería de datos como una etapa del mismo. (Valero, Salvador, & García, 2008). Dentro de ésta se pueden utilizar diversos algoritmos predictivos como:

**Árboles de decisión** C4.5: categorizado como aprendizaje basado en similitudes, los árboles de decisión son uno de los algoritmos más sencillos y fáciles de implementar y a su vez de los más poderosos. Este algoritmo genera un árbol de decisión de forma recursiva al considerar el criterio de la mayor proporción de ganancia de información (gain ratio) (BRITOS & HOSSIAN,

2005), es decir, elige al atributo que mejor clasifica a los datos. Técnica de los k vecinos más cercanos: conocido como algoritmo de aprendizaje basado en instancias, su funcionamiento es muy simple: se almacenan los ejemplos de entrenamiento de datos históricos y cuando se requiere clasificar a un nuevo objeto, se extraen los objetos más parecidos y se usa su clasificación para clasificar al nuevo objeto. (MORALES, 2009) Los vecinos más cercanos a una instancia se obtienen, para el caso de los atributos continuos, utilizando la distancia Euclidiana sobre los n posibles atributos. El resultado de la clasificación por medio de este algoritmo puede ser discreto o continuo. En el caso discreto, el resultado de la clasificación es la clase más común de los k vecinos. (Valero, Salvador, & García, 2008).

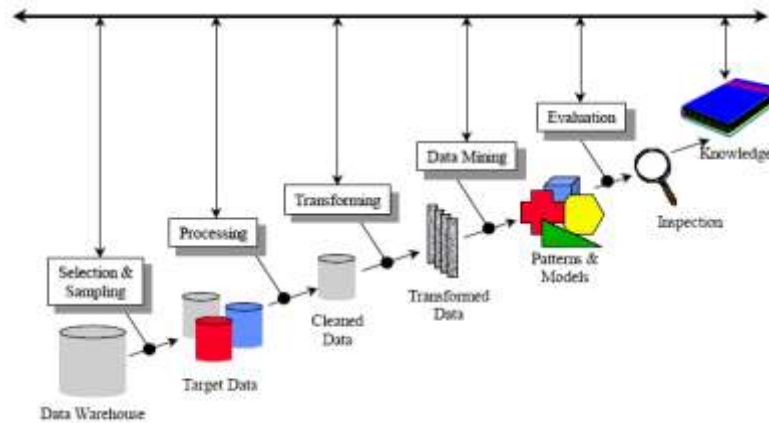
Debido a los grandes volúmenes de datos que una organización puede llegar a tener, el obtener conocimiento a partir de estos no es una tarea fácil. Con este fin investigadores han dado origen a campos de investigación como el descubrimiento de conocimiento en bases de datos (Zhao & Bhowmick, 2006) (Knowledge Discovery in Database - KDD), y en especial, el proceso de minería de datos (Data Mining).

El término 'KDD' es empleado para describir el proceso total de descubrimiento y extracción de conocimiento nuevo, no obvio a partir de un conjunto de datos, el cual está conformado por relaciones y patrones entre los elementos que conforman los datos.

El proceso de KDD abarca varias etapas en su realización, desde la selección de datos que pueden ser necesarios para descubrir conocimiento, hasta visualizar los resultados de dicho descubrimiento. El principal proceso dentro del KDD es la minería de datos 'Data Mining', que es la responsable de buscar, descubrir y extraer el conocimiento desde los datos. (Santamaria, 2015)

La taxonómica general del proceso KDD, como se ve en la figura 2(tomada de Jiawei Han)





Fuente: Eckert & Suénaga, 2015

Según (Eckert & Suénaga, 2015) La minería de datos en la educación (MDE) no es un concepto nuevo, en los últimos años su estudio ha tomado relevancia. La utilización de las técnicas de MD permite descubrir fenómenos dentro del ámbito educativo; de esta forma, es posible determinar la probabilidad de desertar o continuar con sus estudios. (Anand Kumar y Uma, 2009; Ramaswani y Bhaskaran, 2009, Romero y Ventura, 2010).

### **Diferencias entre Minería de Datos y Estadística**

La estadística analiza *muestras* de datos para posteriormente inferir sobre toda la población; la minería de datos trata de usar la totalidad de la población para el análisis (base de datos). Por otro lado, la estadística en un gran porcentaje de casos supone que los datos se comportan de acuerdo a ciertas distribuciones de probabilidad (normal, binomial, geométrica, poisson etc, que llevan a hipótesis que pueden o no llegar a comprobarse), mientras que la minería de datos se basa en técnicas exploratorias o predictivas, que no suponen seguir distribuciones de probabilidad.

### **Diferencias entre la Minería de Datos y el Análisis de Datos (Analyse des Données-Exploratory Data Analysis)**

La minería de datos trata de automatizar la mayor cantidad de procesos, el Análisis de Datos tiene siempre una persona detrás haciendo el análisis manual, por tanto, se dificulta tratar con grandes volúmenes de información.

### **Diferencias entre Minería de Datos y Bodega de datos (Data Warehouse)**

Una Bodega de Datos es un almacén de datos que tiene datos operacionales, agregados (sumarizaciones), datos históricos, evolutivos y en ciertos casos datos externos que están ligados a la empresa, ésta información es accesible a todas las aplicaciones orientadas a la toma de decisiones.

Una bodega de datos es la base de la información para la minería de datos y los dos forman parte del proceso KDD (Proceso de Extracción del Conocimiento-*Knowledge Discovery in Databases*).

### **Diferencias entre Minería de Datos y Machine Learning**

Machine Learning es un área de la Inteligencia Artificial (IA) (IA evolucionada) que trata principalmente sobre la generación de nuevos algoritmos, mediante programación, se enfoca en la elaboración de programas que puedan aprender eficiente y rápidamente; mientras que la minería de datos tiene una mayor influencia de la estadística y se enfoca en la aplicación de los métodos, la interpretación adecuada de los resultados y su influencia directa en la toma de decisiones.

### **Aplicaciones de la Minería de Datos**

- Retención de clientes y Manejo de riesgo (score de crédito – puntuación de crédito)
- Patrones de Compra (algoritmos de recomendación)
- Detección de fraude (detección de transacciones fraudulentas en bancos, tarjetas de crédito, etc.)
- Segmentación de clientes (categorizar los clientes, por ejemplo: dirigir campañas publicitarias por segmentos)
- Predicción de Ventas (Ventas que tendrá un supermercado para evitar el exceso de productos en inventario y no se pierdan lo productos)

### **División de los modelos que usa la Minería de Datos**

- Descriptivo
- Predictivos



Fuente: Eckert & Suénaga, 2015

### **Análisis Descriptivo de Minería de Datos (Modelos de Descripción)**

Búsqueda de patrones en grandes volúmenes de información; son modelos no supervisados. Entre las técnicas más usadas en este método se tiene:

- OLAP (Visualización)
- Clustering
- Métodos factoriales – Métodos de reducción de la dimensión como el Análisis en componentes principales (ACP), etc.

### **Clustering**

Su función principal es encontrar grupos de datos.

### **Análisis de Predicción de Minería de Datos (Modelo de Predicción)**

Los modelos de predicción son abstracciones de representaciones del mundo real aplicadas al tratamiento predictivo. (Tutiempo Network, 2015).

Los modelos de predicción se aplican generalmente cuando la variable a estudiada es cualitativa, se usa generalmente para pronosticar futuros fraudes, deserciones, score crediticio, etc.

Son modelos supervisados, las técnicas más usadas para el análisis predictivo en Minería de datos son:

- Series de tiempo

- Análisis Discriminante
- Regresión
- Árboles de decisión
- Métodos de potenciación, etc.

### Datos sobre los que trabaja la minería de datos

Hoy en día los datos se encuentran almacenados en:

- Base de datos relacionales
- Bodegas de datos
- Base de datos transaccionales
- Bases de datos orientada a objetos y simbólicas
- Bases de datos espaciales sistemas de información geográfica
- Series cronológicas de datos y datos temporales
- Bases de datos de texto
- Bases de datos multimedia
- www (Web Mining)

La Minería de Datos se trabaja sobre todo tipo de información, sin embargo, es conveniente concentrar los datos en una matriz, para facilidad

### Minería de Datos y Business Intelligence



Fuente: Eckert & Suénaga, 2015

### **Que no es Minería de datos**

La minería de datos NO es un modelo determinístico (modelo matemático); se acerca más a un modelo probabilístico, ya que, al hacer una predicción, lo que en realidad se hace es dar un valor de probabilidad de un evento.

### **Factores Socio demográficos**

- Un tamaño de la familia mayor a los promedios nacional, rural o urbano.
- El padre o la madre no tienen presencia directa en los hogares.
- Nivel de instrucción de los padres y las madres es muy bajo.
- Salarios que reciben los miembros de los hogares que trabajan son muy bajos.
- Ingresos del hogar son limitados en relación al costo estimado para la canasta básica.
- Jefe del hogar no es el padre o la madre del desertor o
- Embarazos.
- Compromisos laborales que impiden atender adecuadamente a la propia educación.
- Percepción alta de los problemas económicos (INEC, 2008)

### **Indicadores**

Datos que nos permiten medir de forma objetiva los sucesos para poder respaldar acciones.

Son variables que intentan medir u objetivar en forma cuantitativa o cualitativa, sucesos colectivos para así, poder respaldar acciones. (CORREA, 2012).

### **Atributos de un buen indicador**

**Disponibilidad:** los datos básicos para la construcción del indicador deben ser de fácil obtención sin restricciones de ningún tipo.

**Simplicidad:** el indicador debe ser de fácil elaboración.

Validez: la validez de los indicadores significa que éstos deben tener la capacidad de medir realmente el fenómeno que se quiere medir y no otros.

Especificidad: si un indicador no mide realmente lo que se desea medir, su valor es limitado, pues no permite la verdadera evaluación de la situación, al reflejar características que pertenecen a otro fenómeno paralelo.

**Confiabilidad**: los datos utilizados para la construcción del indicador deben ser fidedignos (fuentes de información satisfactorias).

**Sensibilidad**: el indicador debe ser capaz de poder identificar las distintas situaciones aún en áreas con distintas particularidades, independientemente de la magnitud que ellas tengan en la comunidad.

**Alcance**: el indicador debe sintetizar el mayor número posible de condiciones o de distintos factores que afectan la situación descrita por dicho indicador. En lo posible el indicador debe ser globalizador. (Correa, 2012)

### **Indicadores**

Actividad Académica de los estudiantes

Estudiantes graduados

Estudiantes egresados (Es desertor)

Estudiantes inactivos por deserción voluntaria

Deserción por reglamento (pérdidas)

### **Trayectoria académica**

Indudablemente la educación superior constituye la puerta de acceso más importante a la sociedad del conocimiento, toda vez que representa el medio ideal para el acrecentamiento del capital humano y de su inteligencia individual y colectiva; así fue reconocido en la Conferencia Mundial sobre Educación Superior organizada por la UNESCO en 1998: "...la educación superior y la investigación forman hoy en día la parte fundamental del desarrollo cultural, socioeconómico y ecológicamente sostenible de los individuos, las comunidades y las naciones" (UNESCO, 1998).

Según (Chavez & Zurita, 2017) sobre la Trayectoria Académica menciona como variables, Tipo de Colegio de donde provienen, especialización en el Colegio.

## **Deserción Estudiantil**

Abandono de la institución educativa. Se trata de aquella situación en la que el alumno después de un proceso acumulativo de separación o retiro, finalmente, comienza a retirarse antes de la edad establecida por el sistema educativo sin obtener un certificado.

### **2.5. Hipótesis**

Las técnicas de predicción incidirán en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato.

### **2.6. Señalamiento de Variables**

#### **2.6.1. Variable Independiente**

- Técnicas de Predicción

#### **2.6.2. Variable Dependiente**

- Deserción Estudiantil

## **CAPÍTULO III**

### **METODOLOGÍA**

#### **3.1. Enfoque de la Investigación**

tomó en cuenta la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas para analizar la Deserción Estudiantil, se recolecto datos de los estudiantes de los libros de matrículas y fichas estudiantiles, basándonos en datos numéricos por semestre o número de créditos que los estudiantes toman en la actualidad, se encuesta a los estudiantes que se encuentran actualmente en la Carrera de Docencia en Informática, sobre cuáles serían los posibles motivos que provocarían una deserción, a los docentes, y estudiantes desertores para la determinación de indicadores.

Por lo cual se utilizó variables cualitativas para la determinación de las causas y cuantitativas para los datos numéricos como la edad, promedio, los datos estudiados fueron tomados a partir del año 2006.

Para el análisis e interpretación de los resultados, en las preguntas de selección múltiple se aplicó la moda para determinar la frecuencia mayor.

#### **3.2. Modalidad Básica de la Investigación**

##### **Investigación de Campo**

La investigación de campo se llevó a cabo mediante la observación directa de los hechos en lugares como: aula de clases, secretarías, y diferentes contextos con los estudiantes activos y desertores de la Carrera de Docencia en Informática, por ello se obtuvo información de la fuente, en su contexto original.



### **Documental o Bibliográfica**

Se revisaron artículos científicos de las revistas virtuales Scopus y Scielo de donde se tomó información relacionada con las variables de estudio, estos datos sirvieron como base para la realización del modelo de predicción, de igual manera se accedió a información de Google académico para la lectura de texto online.

### **3.3. Niveles o Tipos de Investigación**

#### **3.3.1. Exploratorio**

La base de la investigación fue de tipo exploratorio se llevó a cabo mediante la toma de datos históricos de los estudiantes, así como investigaciones realizadas, en otros contextos.

el inicio de la investigación se tomó en cuenta los datos de estudiantes desertores, y no desertores, así como de diversos docentes.

#### **3.3.2. Descriptivo**

Se describieron las causas que propician la Deserción Estudiantil se procedió a la identificación de las características socios demográficas, formas de conducta, actitudes de los estudiantes, para establecer patrones de comportamientos concretos sobre una posible deserción.

#### **3.3.3. Correlacional**

Se determinó la relación que existe entre las variables propuestas para la investigación con la ayuda de métodos estadísticos, así pues, se puede mencionar.

### **3.4. Población y muestra**

**Población:** Se trabajará con la siguiente población:

**Tabla 1: Población**

Estudiantes Carrera de Docencia en Informática	70
Estudiantes Carrera de Docencia en Informática a partir 2006 (tomados en cuenta para la minería de datos)	485
Docentes	10
Estudiantes egresados	30
Estudiantes desertores	5
<b>Total Población</b>	<b>600</b>

Elaborado por: Blanca Cuji

## **Muestra**

Se trabajó con la totalidad de la población propuesta, por ser relativamente pequeña

### 3.5. Operacionalización de Variables

Tabla 2: Operacionalización variable independiente

Variable Independiente: Técnicas de Predicción				
CONCEPTUALIZACIÓN	DIMENSIONES	INDICADORES	ITEMS	TÉCNICA INSTRUMENTO
Una de las técnicas de predicción utilizadas comúnmente es la minería de Datos: Procedimiento, reglas, normas o protocolos que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos. El objetivo general del proceso de minería de datos consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior	Base de Datos	Archivos planos	¿Cuál de las siguientes razones propician la Deserción Estudiantil?	TECNICA Encuesta INSTRUMENTO Cuestionario
		Relacionales	Dificultades económicas	
		Multidimensionales	Sanción Disciplinaria	
			Aplazamiento de Semestre	
			Bajo rendimiento académico	
			Cambio de carrera	
			Cambio de ciudad	
			Dificultades familiares	
			Enfermedad	
		Estadísticos	Ubicación laboral	
	Indicadores	Económicas	Otros	
		Sociales	La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos:	
			Humanos	
			Materiales	
			Tecnológicos	
			¿Otros Cual? .....	
			¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?	

---

Si ( ) No ( )

Al culminar el semestre octubre 2015 – marzo 2016, tú crees que:

- Aprobaré todas mis materias con buenas calificaciones.
- Aprobaré todas mis materias calificaciones regulares.
- Probablemente repruebe una materia.
- No creo aprobar este semestre

¿Cuánto tiempo dedicas a tus estudios como parte de tu trabajo autónomo (fuera e clases)?

- Menos de 1 h.
- De 1 a 2 h.
- Más de 2 h.

La decisión de estudiar la carrera universitaria en su momento...

- La tome yo.
- Mis padres y/o familiares influyeron en mi decisión.

¿Tienes alguna discapacidad física o psicológica?

- Sí.
- No.

¿Tienes o has tenido alguna enfermedad grave?

- Sí. ... ¿Cuál?.....
- No.

En cuanto al nivel económico de ingreso en tu familia es...

- Alto.
- Medio.
- Bajo.
- Muy bajo.

¿Dispones de los recursos económicos para costear todos los gastos que se necesitan para estudiar?

- Sí.
- No.

¿Cuentas con algún tipo de ayuda o beca como ayuda para realizar tus estudios?

- Sí.
  - No.
-

**Tabla 3: Operacionalización variable dependiente**

<b>Variable Dependiente: Deserción Estudiantil</b>				
<b>CONCEPTUALIZACIÓN</b>	<b>DIMENSIONES</b>	<b>INDICADORES</b>	<b>ITEMS</b>	<b>TÉCNICA INSTRUMENTO</b>
Abandono escolar es aquella situación en la que el alumno después de un proceso acumulativo de separación o retiro, finalmente, comienza a retirarse antes de la edad establecida por el sistema educativo sin obtener un certificado	Abandono Escolar	Separación Retiro	¿Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia? 10% 20% 30% 40% 50% 60% 70% 80% No conoce	TECNICA Encuesta INSTRUMENTO Cuestionario
	Sistema Educativo	Créditos		
	Factores	Económicos		
		Sociales Demográficos		
			Los estudiantes abandonan sus estudios cuando cursan el..... de sus créditos 10% 20% 30% 40% 50% 60% 70% 80% No conoce	
			Del nivel de estudios de tus padres. Marca con una X: MADRE    PADRE Primaria    Primaria Secundaria    Secundaria Superior    Superior Postgrado    Postgrado No sé    No sé	
			¿En tu casa cuentas con un espacio adecuado (con mesa o escritorio) para estudiar? <input type="checkbox"/> Sí. <input type="checkbox"/> No.	
			¿Tus padres, fomentan y estimulan tus estudios?	

- 
- Sí.  
 No.

¿Consideras que los conocimientos que estas adquiriendo en la institución son de utilidad para la vida?

- Sí.  
 Sólo algunos.  
 No.

¿Es alguna asignatura especialmente difícil para ti?

- No.  
 Sí..... ¿Cuál?

En el examen de admisión de ingreso a la universidad que nota obtuviste:

- Excelente.  
 Muy bueno.  
 Bueno.  
 Regular.  
 Malo.  
 No presenté.

¿Has reprobado alguna asignatura de manera que te hayas retrasado en algún periodo académico?

- Sí.  
 No.

¿Abandonado en alguna ocasión sus estudios en una institución educativa?

Si No

Si su respuesta fue positiva escoja cuál de estos factores influyeron en su decisión

- Económicos
  - Sociales (migración, religión, nivel económico)
  - Demográficos (edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)
  - Rendimiento Académico
- 

Elaborado por: Blanca Cuji

### **3.6. Técnicas e instrumentos de recolección de información**

La técnica usada para la recolección de información fue la encuesta diseñada para recolectar información y opiniones sobre la Deserción Estudiantil y las técnicas predicción, los datos recopilados contribuyeron a la determinación de indicadores sobre las variables planteadas. El instrumento utilizado fue el cuestionario diseñado para Estudiantes, Egresados, Estudiantes desertores, y Docentes de la Carrera de Docencia en Informática, las preguntas se construyeron en base a las dimensiones e indicadores obtenidos en la operacionalización de variables, se diseñaron además encuestas online para la toma de datos de los estudiantes desertores y egresados.

### **3.7. Procesamiento de la Información**

El procesamiento de datos diversas funciones, que se detallan a continuación:

- Se elaboró una validación de los datos por el investigador para asegurar que la información suministrada sea correcta y útil para la investigación
- Se clasificó la información de acuerdo a los conglomerados (Estudiantes, egresados, estudiantes desertores, y docentes).
- Se tabularon los datos, presentándolos en tablas con porcentajes y gráficos
- Se analizó e interpretó la información recolectada
- Se presentó la información de manera detallada y resumida

Todo este proceso sirvió para la producir de información significativa para la investigación.

### **3.8. Validez y Confiabilidad**

#### **3.8.1. Validez**

La validez del cuestionario se realizó mediante el análisis bibliográfico de las variables de la hipótesis, y el aporte del tutor de la investigación, experto en la materia.

Para la valides de la información se realizó un análisis estadístico de datos cuantitativos como la edad, el rendimiento académico etc., mediante la aplicación del coeficiente de correlación, cruzando diferentes variables para medir su grado de correlación.

### **3.8.2. Confiabilidad**

Para medir la confiabilidad del cuestionario se administró el mismo instrumento de medición, a distintos grupos (Estudiantes, Egresados, Estudiantes desertores, Docentes), esto permitió medir la consistencia y coherencia de la información mediante la visualización de la escasa variación de los patrones de respuesta entre los grupos aplicados.

### **3.9. Plan de recolección de la información**

La información se recopiló mediante el uso de instrumentos de investigación, como encuestas aplicadas a los estudiantes, docentes, estudiantes desertores y egresados de la Carrera de Docencia en Informática.

Se recogió la información utilizando cuestionarios, mediante una revisión de la información de manera sistemática, analítica y reflexiva.

### **3.10. Plan de procesamiento de la información**

El análisis y procesamiento de la información se realizó conforme el siguiente detalle:

- Recolección y análisis de la información mediante un estudio crítico para su posterior depuración.
- Se tabularon y procesaron los datos obtenidos en las encuestas mediante el uso de gráficos apoyados en aplicaciones ofimáticas como hojas de cálculo, editores de texto, etc.



- Se procedió al análisis e interpretación de los datos con base en el marco teórico propuesto.
- Comprobación de la Hipótesis mediante la aplicación del método estadístico apropiado.
- Determinación de las conclusiones y recomendaciones
- Propuesta de solución al problema detectado.

## CAPÍTULO IV

### ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

#### 4.1. Análisis de Resultados

##### Encuesta realizada a estudiantes de la Carrera de Docencia en Informática

##### I. DATOS GENERALES

Tabla 4: Edad Estudiantes

Edad	Frecuencia	Porcentaje
16-20	14	20%
21-25	44	63%
26-35	10	14%
Más de 36	2	3%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

##### Análisis e interpretación

De los 70 estudiantes de la Carrera de Docencia, 44 tienen edades de 21 a 25 años, esto corresponde al 63% de la población, el 20% de estudiantes tienen edades entre 16 a 20 años; 12 estudiantes tienen edades de 26 a 35 y solo el 3% de la población supera los 36 años.

Se puede evidenciar que se forman dos grupos de edades, el primero corresponde a la masa poblacional centrada en edades de 16 a 25 años con más de la mitad de la población y el segundo grupo de los estudiantes de más de 26 años.

**Tabla 5: Lugar de nacimiento-Estudiantes**

<b>Lugar de nacimiento:</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Ambato	53	76%
Salcedo	1	1%
Zamora Chinchipe	3	4%
Salasaca	1	1%
Puyo	2	3%
Pillaro	1	1%
Latacunga	2	3%
El Empalme	1	1%
Cuenca	1	1%
Pelileo	1	1%
Guayaquil	1	1%
Babahoyo	1	1%
Quito	2	3%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 76% de los estudiantes de la Carrera de Docencia en Informática tienen como lugar de nacimiento a la ciudad de Ambato, en menor cantidad se encuentran estudiantes de Zamora Chinchipe, Puyo, Latacunga, Cuenca, Guayaquil, Babahoyo, Quito, así como de cantones de la provincia de Tungurahua como Pillaro, y Pelileo.

**Tabla 6: Genero- Estudiantes**

<b>Género</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Hombre	41	59%
Mujer	29	41%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El 59% de la población estudiantil de la Carrera de Docencia en Informática son de género masculino; y el 41% restante son de género femenino.

**Tabla 7: Estado Civil**

<b>Estado Civil</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Casado	13	19%
Soltero	54	77%
Unión Libre	3	4%
Viudo	0	0%
Divorciado	0	0%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El 77% de la población es soltera, mientras que el 19% de los estudiantes son de estado civil casado y el 4%, unión libre.

**Tabla 8: Discapacidad**

<b>Capacidad especial</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	5	7%
No	65	93%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

<b>Tipo de Discapacidad</b>	<b>frecuencia</b>
Visual	3
Físico:	1
Psicológico:	1
Intelectual:	0
Sensorial:	0
<b>Total</b>	<b>5</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 93% de la población no posee una discapacidad, mientras que el 5% restante tiene discapacidad de tipo visual, física y psicológica.

**Tabla 9: Identificación Étnica**

Identificación Étnica	Frecuencia	Porcentaje
Mestizo	62	89%
Blanco	2	3%
Indígena	6	8%
Afro descendiente	0	0%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

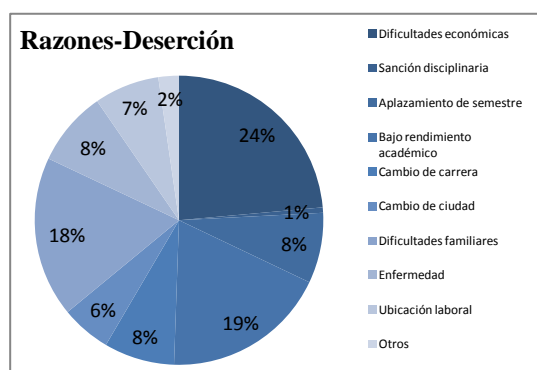
## Análisis e interpretación

89% de los estudiantes se identifica como mestiza, en menor porcentaje se encuentra la población indígena y blanca.

## II. CUESTIONARIO

### 1. ¿Cuál de las siguientes razones propician la Deserción Estudiantil?

(pregunta de selección múltiple)



**Tabla 10: Razones que propician la Deserción Estudiantil**

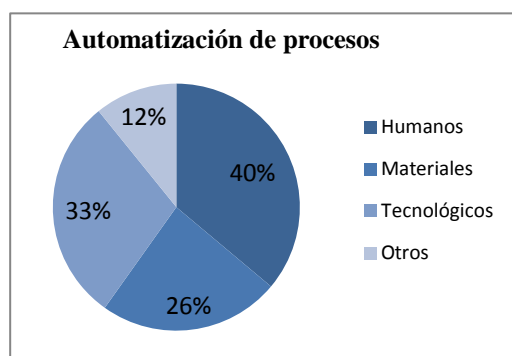
Opciones	Frecuencia	Porcentaje
Dificultades económicas	42	24%
Sanción disciplinaria	1	1%
Aplazamiento de semestre	14	8%
Bajo rendimiento académico	33	19%
Cambio de carrera	14	8%
Cambio de ciudad	10	6%
Dificultades familiares	32	18%
Enfermedad	15	8%
Ubicación laboral	13	7%
Otros	4	2%
<b>Total</b>	<b>178</b>	<b>24%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

Según el punto de vista de los estudiantes las dificultades económicas y familiares, así como el bajo rendimiento propician la Deserción Estudiantil, por otra parte, factores como sanción disciplinaria y cambio de ciudad tienen menor incidencia en el abandono estudiantil.

### **2. La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos (pregunta de selección múltiple)**



**Tabla 11: Automatización de procesos**

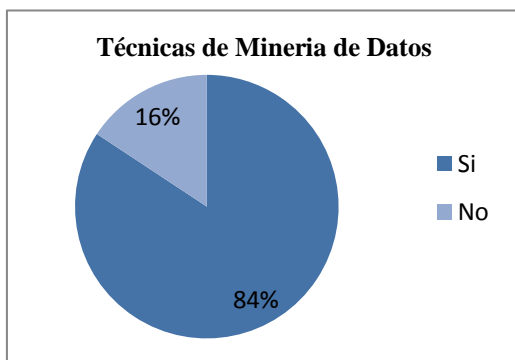
Opciones	Frecuencia	Porcentaje
Humanos	32	40%
Materiales	21	26%
Tecnológicos	26	33%
Otros	1	12%
<b>Total</b>	<b>80</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

Mediante la automatización de procesos sobre Deserción Estudiantil se optimizará el uso de recurso humanos según la moda encontrada en el procesamiento de los datos, en menor cantidad se considera a los recursos tecnológicos y materiales.

### **3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?**



**Tabla 12: Técnicas de Predicción**

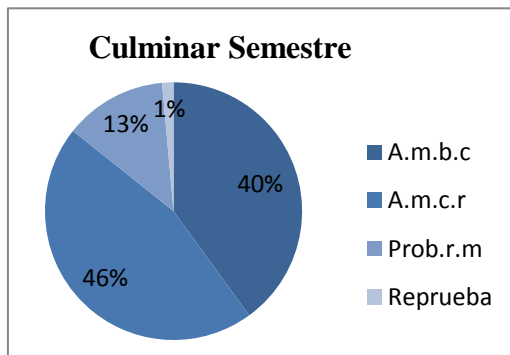
Opciones	Frecuencia	Porcentaje
Si	59	84%
No	11	16%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 84% de estudiantes considera que las técnicas de predicción, basada en minería de datos ayudarán a detectar patrones de Deserción Estudiantil en la Carrera.

### 4. Al culminar el semestre Octubre 2015- Marzo 2016, usted cree que:



Elaborado por: Blanca Cuji

Tabla 13: Culminar Semestre

Opciones	Frecuencia	Porcentaje
Aprobará todas sus materias con buenas calificaciones.	28	40%
Aprobará todas sus materias con calificaciones regulares	32	46%
Probablemente reprueben alguna materia	9	13%
No creo que apruebe este semestre	1	1%
<b>Total</b>	<b>70</b>	<b>100%</b>

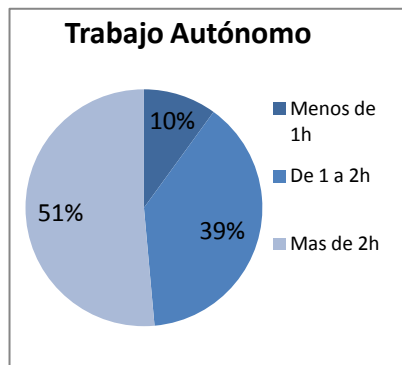
Elaborado por: Blanca Cuji

## Análisis e interpretación

la percepción del 46% de los estudiantes que aprobará el semestre con calificaciones regulares, en tanto que el 40% considera aprobar sus materias con buenas calificaciones, el 13% sugiere la probabilidad de reprobar alguna materia y el 1% piensa que reprobará el semestre.



**5. Cuanto tiempo dedica a sus estudios como parte de su trabajo autónomo (fuera de clases)?**



**Tabla 14: Dedicación trabajo autónomo**

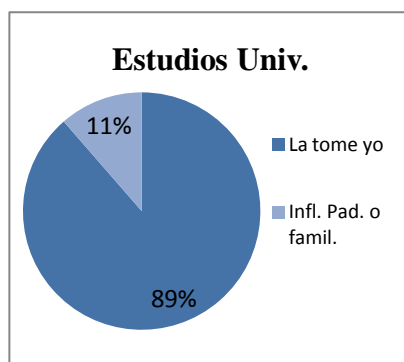
Opciones	Frecuencia	Porcentaje
Menos de 1h	7	10%
De 1 a 2h	27	39%
Más de 2h	36	51%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

Los resultados muestran que el 51% de la población dedica más de dos horas a su trabajo autónomo, el 39% de una a dos horas y el 10% de estudiantes dedica menos de una hora.

**6. La decisión de estudiar la carrera universitaria en su momento fue:**



**Tabla 15: Decisión de tomar una carrera universitaria**

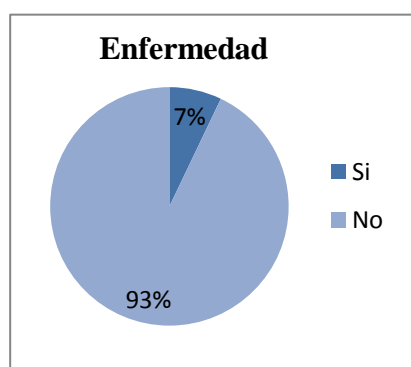
Opciones	Frecuencia	Porcentaje
La tome yo	62	89%
Mis padres y/o familiares influyeron en mi decisión	8	11%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

La decisión de estudiar la Carrera de Docencia en Informática fue tomada en su gran mayoría por el propio estudiante así lo indica la estadística en donde el 89% de los estudiantes expresa su decisión.

### **7. ¿Ha tenido alguna enfermedad grave?**



**Tabla 16: Enfermedad**

Opciones	Frecuencia	Porcentaje
Si	5	7%
No	65	93%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

---

## Enfermedad

Angiofibroma  
y otros

Elaborado por: Blanca Cuji

### Análisis e interpretación

El 93% de los estudiantes no posee algún tipo de enfermedad, el 5% restante sufre de algún tipo de enfermedad

### 8. En cuanto al nivel económico de ingreso en su familia es:

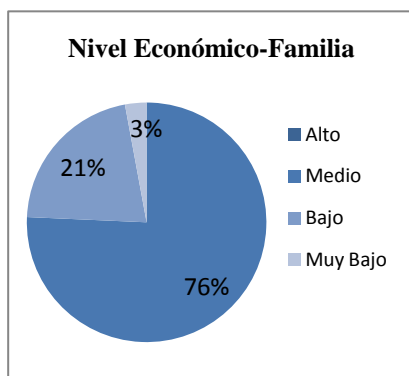


Tabla 17: Nivel Económico

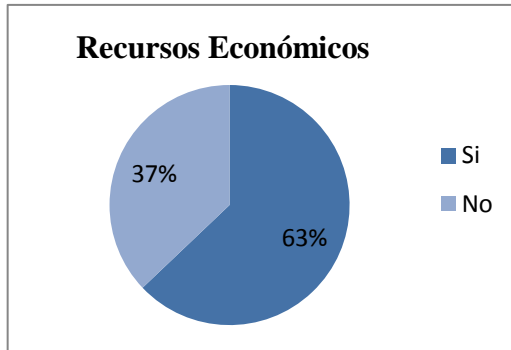
Opciones	Frecuencia	Porcentaje
Alto	0	0%
Medio	53	76%
Bajo	15	21%
Muy Bajo	2	3%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

La mayor parte de población estudiantil de la Carrera tiene un nivel económico medio, seguido del 21% con un nivel bajo y el 3% con muy bajo.

**9. ¿Dispone de los recursos económicos para costear todos los gastos que se necesita para estudiar?**



**Tabla 18: Gastos de estudio**

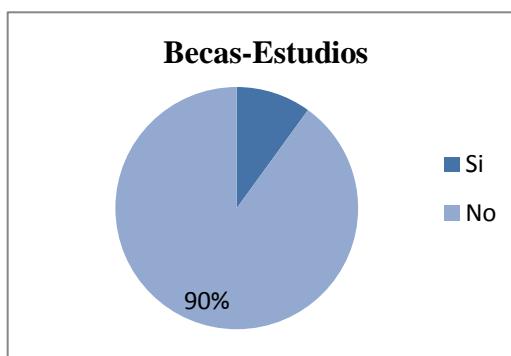
Opciones	Frecuencia	Porcentaje
Si	44	63%
No	26	37%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

El 37% de la población que corresponde a 26 estudiantes no disponen de recursos económicos para costear los gastos de sus estudios.

**10. ¿Cuenta con algún tipo de ayuda o beca para realizar sus estudios?**



**Tabla 19: Ayuda para realizar estudios**

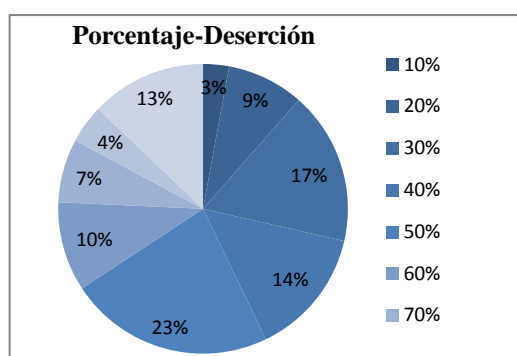
Opciones	Frecuencia	Porcentaje
Si	7	10%
No	63	90%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 90% de los estudiantes, solventa sus estudios con recursos propios, solo el 10% recibe ayuda económica para sus estudios.

### **11. Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia**



**Tabla 20: Porcentaje de deserción**

Opciones	Frecuencia	Porcentaje
10%	2	3%
20%	6	9%
30%	12	17%
40%	10	14%
50%	16	23%
60%	7	10%
70%	5	7%
80%	3	4%
No conoce	9	13%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 23% de la población considera que existe un 50% de Deserción Estudiantil según la moda más representativa de los datos, por otro lado, se encuentran porcentajes bajos como 30%, 40%, 60% y 20%, sin embargo, existe un porcentaje del 13% de estudiantes que no conocen sobre el tema.

### 12. Los estudiantes abandonan sus estudios cuando cursan el...de sus créditos.

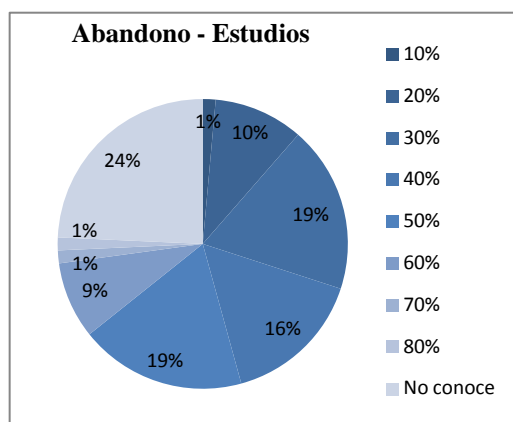


Tabla 21: Nivel de abandono de estudios

Opciones	Frecuencia	Porcentaje
10%	1	1%
20%	7	10%
30%	13	19%
40%	11	16%
50%	13	19%
60%	6	9%
70%	1	1%
80%	1	1%
No conoce	17	24%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

Existe un desconocimiento por parte de la población encuestada sobre el número de créditos que logran alcanzar los estudiantes desertores antes de abandonar sus estudios, como se evidencia en los datos tomados, en donde la moda más representativa es de 24% que corresponde al “No conoce”, sin embargo porcentajes mínimos de la población considera que los estudiantes abandonan sus estudios al cumplir el 30%, 40% y 50% de sus créditos.

### 13. Del nivel de estudio de sus padres. Marca con una x.

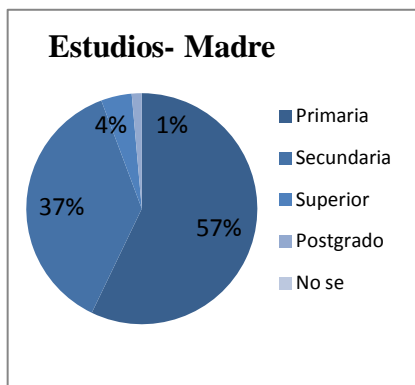
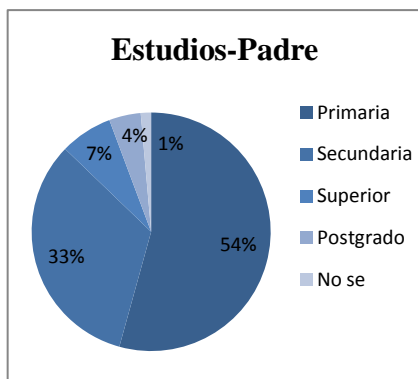


Tabla 22: Nivel de estudios de los padres

Opciones	Frecuencia	Porcentaje
<b>MADRE:</b>		
Primaria	40	57%
Secundaria	26	37%
Superior	3	4%
Postgrado	1	1%
No se	0	0%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji



**Tabla 23: Nivel de estudio del padre**

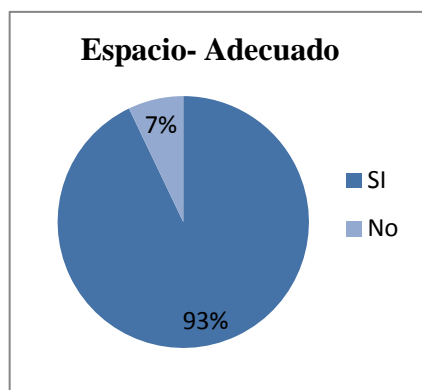
<b>PADRE:</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Primaria	38	54,3%
Secundaria	23	32,9%
Superior	5	7,1%
Postgrado	3	4,3%
No se	1	1,4%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El nivel de estudio del padre y la madre de la mayoría de estudiantes es primaria y secundaria, muy pocos padres tienen un nivel de estudios superior o posgrado.

#### **14. ¿En su casa cuenta con espacio adecuado (Con mesa o escritorio) para estudiar?**



**Tabla 24: Espacios para estudiar**

<b>Opciones</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	65	93%
No	5	7%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji



### Análisis e interpretación

En cuanto al espacio físico el 93% de la población encuestada manifiesta que cuentan con una mesa o escritorio, para sus labores escolares.

#### 15. ¿Sus padres, fomentan y estimulan sus estudios?

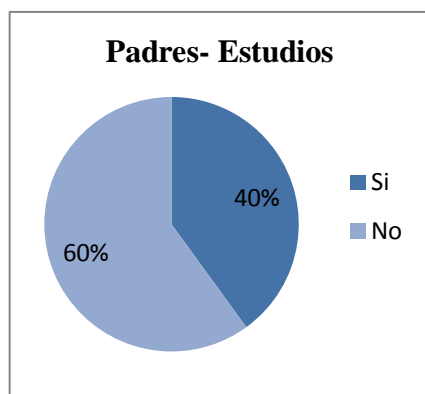


Tabla 25: Estimulo para la realización de estudios

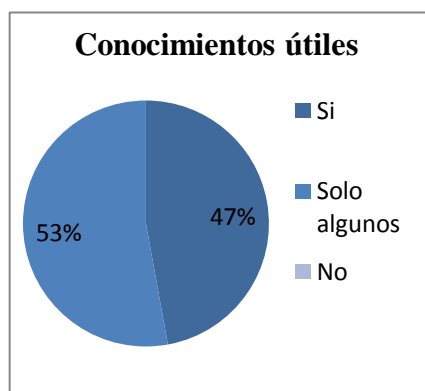
Opciones	Frecuencia	Porcentaje
Si	66	94%
No	4	6%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El 94% de los estudiantes consideran recibir el apoyo y estímulo de sus padres para realizar sus estudios.

#### 16. ¿Considera que los conocimientos que está adquiriendo en la institución son de utilidad para la vida?



**Tabla 26: Conocimientos útiles**

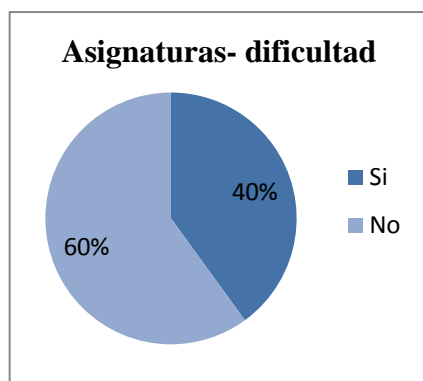
Opciones	Frecuencia	Porcentaje
Si	33	47%
Solo Algunos	37	53%
No	0	0,00%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

Solo algunos de los contenidos impartidos, es considerado de utilidad por los estudiantes, así lo considera el 53% de la población encuestada.

### **17. ¿Es alguna asignatura especialmente difícil para Usted?**



**Tabla 27: Dificultad den las asignaturas**

Opciones	Frecuencia	Porcentaje
Si	28	40%
No	42	60%
<b>Total</b>	<b>70</b>	<b>10%</b>

Elaborado por: Blanca Cuji

---

**¿Cuál?** Base de Datos  
 Programación  
 Ingles  
 Diseño de proyectos  
 Física

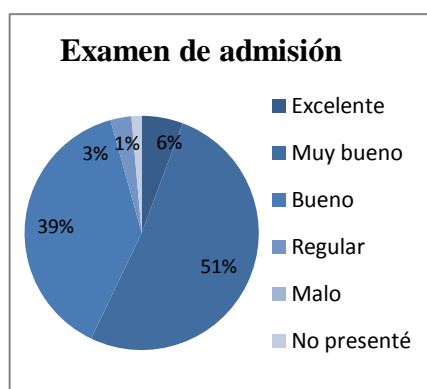
---

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 40% de la población tiene dificultades en materias como programación, base de datos y física, éstas son consideradas difíciles para los estudiantes.

### **18. En el examen de admisión de ingreso a la universidad que nota obtuvo**



**Tabla 28: Examen de admisión**

Opciones	Frecuencia	Porcentaje
Excelente	4	6%
Muy bueno	36	51%
Bueno	27	39%
Regular	2	3%
Malo	0	0%
No presenté	1	1%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 51% de los estudiantes obtuvieron “muy buena” calificación en el examen de ingreso a la Universidad; el 39% una calificación “buena”, el 3% lograron una calificación regular.

### 19. ¿Ha reprobado alguna asignatura de manera que se haya retrasado en algún periodo académico?

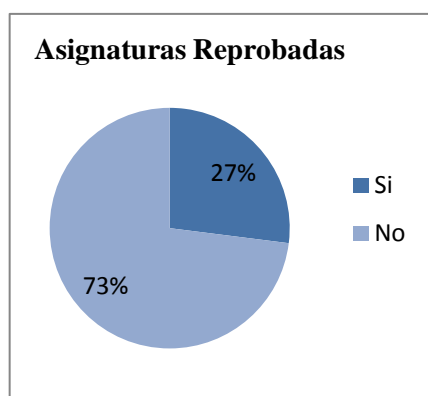


Tabla 29: Reprobación de asignaturas

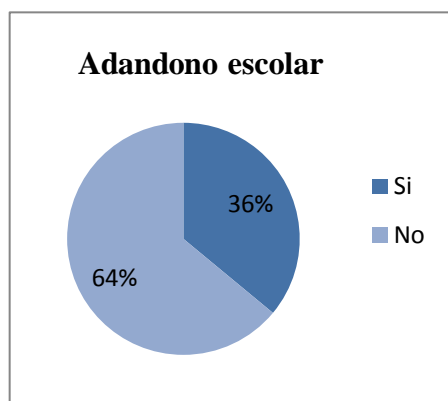
Opciones	Frecuencia	Porcentaje
Si	19	27%
No	51	73%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 73% de los estudiantes ha tenido por lo menos una asignatura retrasada en un periodo académico, solo el 27% del total de la población no ha tenido retrasos en sus estudios.

**20.¿Abandonado en alguna ocasión sus estudios en una institución educativa?**



**Tabla 30: Abandono de estudios**

Opciones	Frecuencia	Porcentaje
Si	25	36%
No	45	64%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Tabla 30: Abandono de estudios**

Factores que influyeron en su decisión	Frecuencia
Económicos	15
Sociales (Migración, religión)	5
Demográfico (Edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)	4
Rendimiento Académico	11

Elaborado por: Blanca Cuji

**Análisis e interpretación**

El 64% de la población manifiesta haber abandonado en alguna ocasión sus estudios. Los factores más relevantes para una posible deserción, son económicos y rendimiento académico.

## Encuesta realizada a estudiantes desertores de la Carrera de Docencia en Informática

### I. DATOS GENERALES

**Tabla 31: Edad - Estudiantes desertores**

Edad	Frecuencia	Porcentaje
16-20	2	40%
21-25	3	60%
26-35	0	0%
Más de 36	0	0%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

Los estudiantes desertores están en una edad de 16 a 25 años.

**Tabla 32: Lugar de nacimiento- Estudiantes desertores**

Lugar de nacimiento	Frecuencia	Porcentaje
Ambato	4	80%
Mocha	0	0%
Guaranda	1	20%
Riobamba	0	0%
Pujili	0	0%
Pasaje	0	0%
Pillaro	0	0%
Pelileo	0	0%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

La mayoría de los estudiantes desertores de la Carrera de Docencia en Informática nacieron en la ciudad de Ambato y Guaranda.

**Tabla 33: Género**

<b>Género:</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Hombre	2	40%
Mujer	3	60%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

Existe mayor porcentaje de deserción en mujeres que en hombres

**Tabla 34: Estado Civil- Estudiantes desertores**

<b>Estado Civil</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Casado	2	40%
Soltero	2	40%
Unión Libre	0	0%
Viudo	0	0%
Divorciado	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El estado civil de los estudiantes desertores es: solteros y casados en mayor porcentaje y divorciados en menor porcentaje.

**Tabla 35: Capacidades especiales- Estudiantes desertores**

<b>Capacidades especiales</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	0	0
No	5	100%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

Ninguno de los estudiantes desertores presenta una capacidad especial(discapacidad).

**Tabla 36: Etnia**

<b>Identificación Étnica</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Mestizo	5	100%
Blanco	0	0
Indígena	0	0
Afrodescendiente	0	0
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

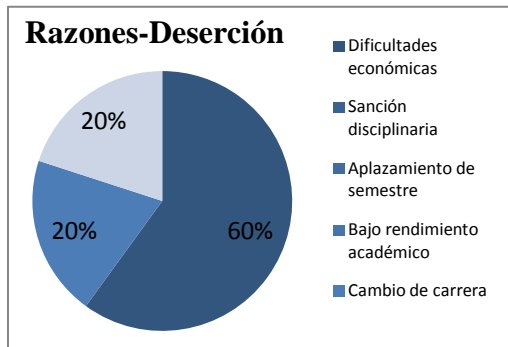
### **Análisis e interpretación**

Todos los estudiantes desertores se identifican como mestizos.



## II. CUESTIONARIO

### 1. ¿Cuál de las siguientes razones propician la Deserción Estudiantil?



**Tabla 37: Razones para la Deserción Estudiantil -Estudiantes Desertores**

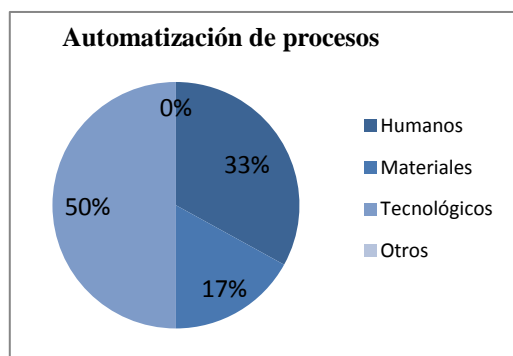
Razones que propician la Deserción Estudiantil	Frecuencia	Porcentaje
Dificultades económicas	3	60%
Sanción disciplinaria	0	0
Aplazamiento de semestre	0	0
Bajo rendimiento académico	0	0
Cambio de carrera	1	20%
Cambio de ciudad	0	0
Dificultades familiares	0	0
Enfermedad	0	0
Ubicación laboral	0	0
Otros	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

Las dificultades económicas, cambios de carrera y otros factores propician la Deserción Estudiantil.

**2. ¿La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos?**



**Tabla 38: Automatización de procesos-Estudiantes Desertores**

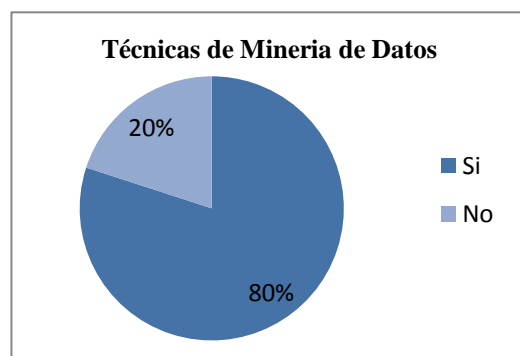
<b>Automatización de procesos en estudios sobre Deserción Estudiantil</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Humanos	2	33%
Materiales	1	17%
Tecnológicos	3	50%
Otros	0	0%
<b>Total</b>	<b>6</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

Realizar estudios sobre Deserción Estudiantil permitirá optimizar el uso de recursos tecnológicos, humanos y materiales.

**3. ¿La técnica de predicción de minería de datos ayudara a detectar patrones de deserción escolar?**



**Tabla 39: Técnicas de Minería de datos-Estudiantes Desertores**

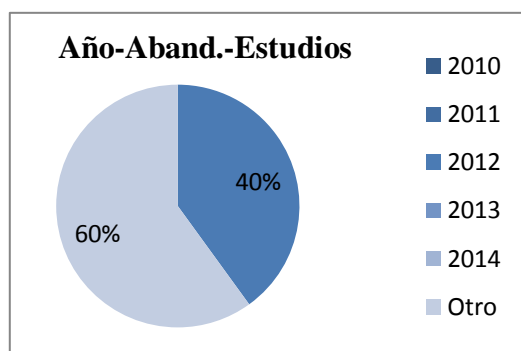
Técnicas de minería de datos	Frecuencia	Porcentaje
Si	4	80%
No	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

Con el uso de técnicas de minería de datos se podrá formar patrones de Deserción Estudiantil que ayuden al estudio sobre este aspecto social, así lo consideran el 80% de los estudiantes desertores, mientras el 20% no lo consideran necesario.

### 4. ¿En qué año abandono sus estudios en la Carrera de Docencia en Informática?



**Tabla 40: Año de deserción-Estudiantes Desertores**

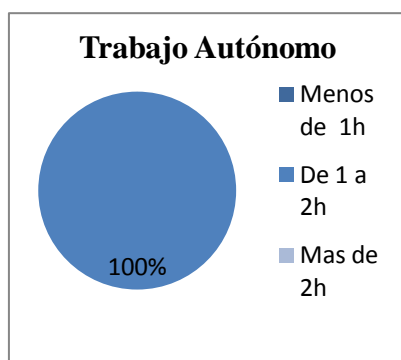
Año abandono sus estudios	Frecuencia	Porcentaje
2010	0	0%
2011	0	0%
2012	2	40%
2013	0	0%
2014	0	0%
Otro	3	60%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El periodo propuesto para el estudio de Deserción Estudiantil desde el año 2010 al 2014, se detecta que los estudiantes desertaron en el año 2012 en un 40% mientras que el 60% en un año diferente al periodo propuesto.

### 5. ¿Cuánto tiempo dedicaba a sus estudios como parte de tu trabajo autónomo (fuera de clases)?



**Tabla 41: Trabajo Autónomo-Estudiantes Desertores**

Trabajo autónomo	Frecuencia	Porcentaje
Menos de 1h	0	0%
De 1 a 2h	5	100%
Más de 2h	0	0%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

La totalidad de estudiantes desertores dedicaban de una a dos horas para sus estudios como parte de su trabajo autónomo.

## 6. La decisión de estudiar la carrera universitaria en su momento...

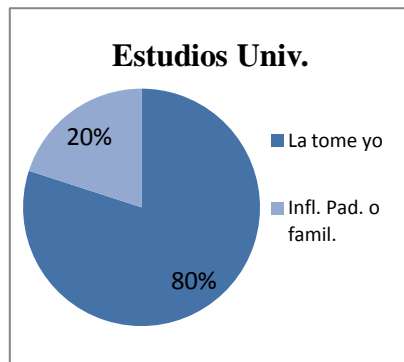


Tabla 42: Carrera Universitaria -Estudiantes Desertores

**Decisión de estudiar la carrera universitaria**

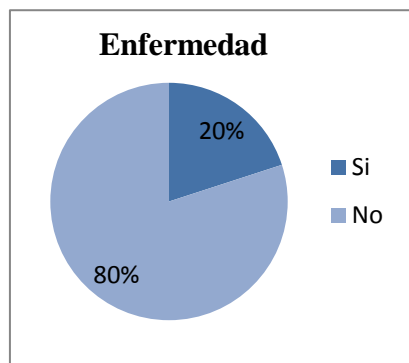
Decisión de estudiar la carrera universitaria	Frecuencia	Porcentaje
La tome yo	4	80%
Mis padres y/o familiares influyeron en mi decisión	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

A su debido tiempo la decisión de estudiar la carrera universitaria fue de los propios estudiantes desertores en un 80%, sin embargo, uno de ellos que corresponden al 20% manifiesta que en su decisión influyeron sus padres.

## 7. ¿Tiene o ha tenido alguna enfermedad grave?



**Tabla 43: Enfermedad-Estudiantes Desertores**

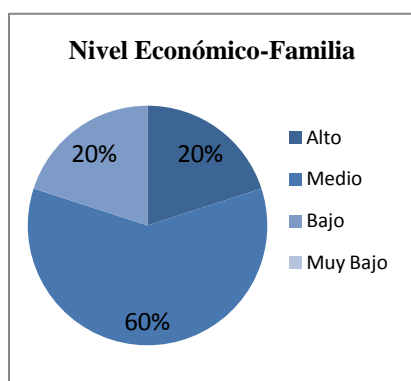
<b>Enfermedad</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	1	20%
No	4	80%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 80% de los estudiantes desertores no han tenido alguna enfermedad, el 20% manifiesta haber padecido un tipo de enfermedad.

### **8. En cuanto al nivel económico de ingreso en su familia es...**



**Tabla 44: Nivel económico de la familia-Estudiantes Desertores**

<b>Nivel económico de su familia</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Alto	1	20%
Medio	3	60%
Bajo	1	20%
Muy Bajo	0	0%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El nivel económico de los estudiantes desertores es medio en un 60%; alto y bajo comparten el 20%.

### 9. ¿Disponía de los recursos económicos para costear todos los gastos que se necesita para estudiar?

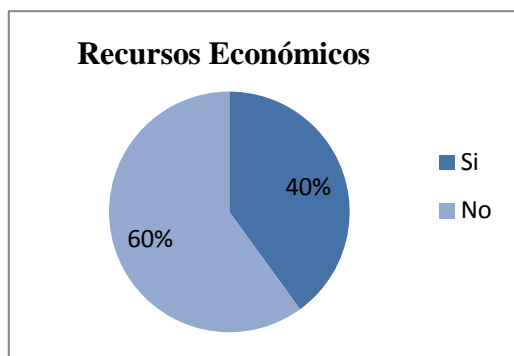


Tabla 45: Gastos Estudiantiles-Estudiantes Desertores

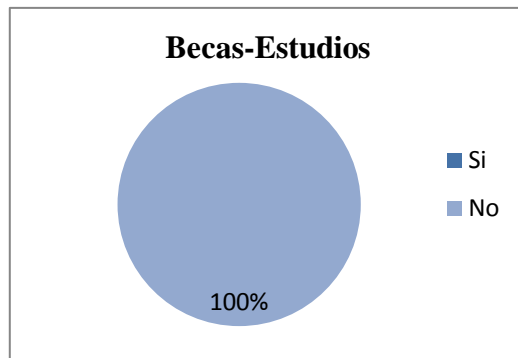
Recursos económicos para estudios	Frecuencia	Porcentaje
Si	2	40%
No	3	60%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El 60% de los estudiantes no disponían de recursos económicos para costear los gastos de su educación; mientras que el 40% manifiesta disponer de los recursos necesarios.

**10. ¿Conto con algún tipo de ayuda o beca para realizar sus estudios?**



**Tabla 46: Becas -Estudiantes Desertores**

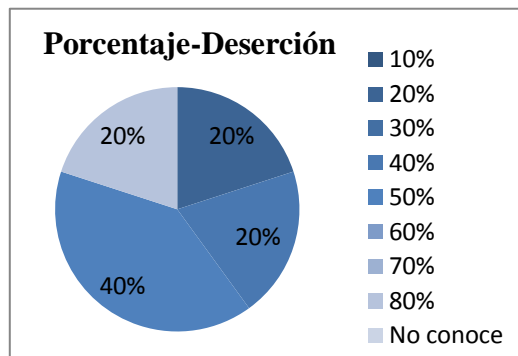
Ayuda o beca para realizar sus estudios	Frecuencia	Porcentaje
Si	0	0%
No	5	100%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

La totalidad de estudiantes desertores no recibieron ningún tipo de ayuda económica para realizar sus estudios.

**11. Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia en Informática?**





**Tabla 47: Porcentaje de deserción -Estudiantes Desertores**

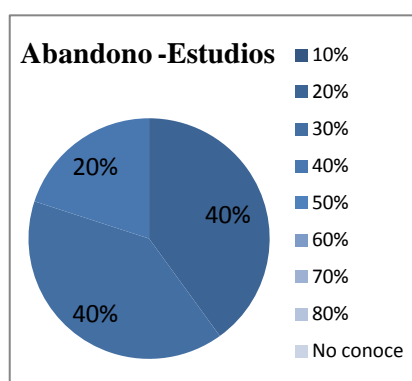
Porcentaje de Deserción Estudiantil	Frecuencia	Porcentaje
10%	0	0
20%	1	20%
30%	0	0
40%	1	20%
50%	2	40%
60%	0	0
70%	0	0
80%	1	20%
No conoce	0	0
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 40% de los estudiantes desertores considera que hay un 50% de Deserción Estudiantil en la Carrera de Docencia en Informática y el 20% de estudiantes toma en consideración porcentajes de 20%, 40% y 80%.

### **12. Los estudiantes abandonan sus estudios cuando cursan el... de sus créditos.**



**Tabla 48: Créditos tomado -Estudiantes Desertores**

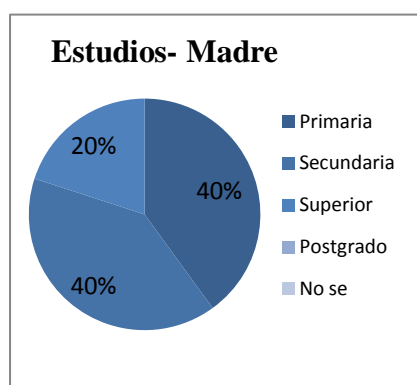
<b>Abandono de estudios</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
10%	0	0
20%	2	40%
30%	2	40%
40%	1	20%
50%	0	0
60%	0	0
70%	0	0
80%	0	0
No conoce	0	0
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

De la totalidad de estudiantes encuestados el 40% considera que los estudiantes abandonan sus estudios aproximadamente cuando cumplen el 20% o 30% de sus estudios; por otro lado, el 20% considera que la deserción se da después de haber cumplido el 40 % de sus créditos.

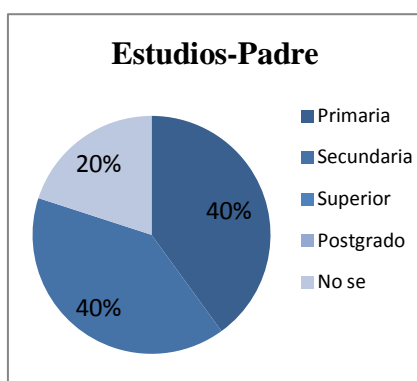
### **13. Del nivel de estudios de sus padres. Marca con una X:**



**Tabla 49: Estudios de los padres-Madre**

<b>Estudio de los padres</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>MADRE:</b>		
Primaria	2	40%
Secundaria	2	40%
Superior	1	20%
Postgrado	0	0
No se	0	0
<b>Total</b>	<b>5</b>	<b>100</b>

Elaborado por: Blanca Cuji



**Tabla 50: Estudios de los padres-Padre**

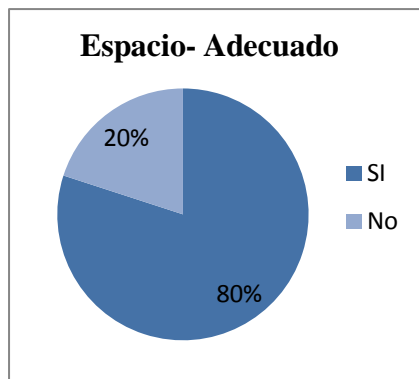
<b>PADRE:</b>		
Primaria	2	40%
Secundaria	2	40%
Superior	0	0
Postgrado	0	0
No se	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El nivel de estudio de los padres en su gran mayoría es de primaria y secundaria en el caso del padre y la madre.

**14. ¿En su casa contaba con un espacio adecuado (con mesa o escritorio) para estudiar?**



**Tabla 51: Espacio adecuado-Estudiantes Desertores**

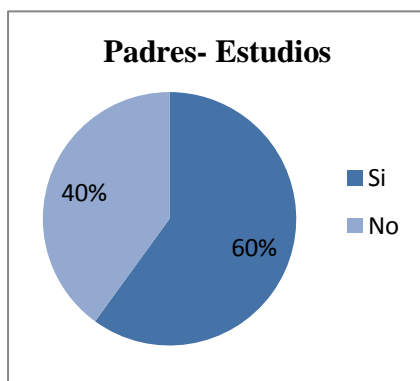
<b>Espacio adecuado</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
SI	4	80%
No	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

El 80% de los estudiantes manifiesta que su espacio para estudiar era adecuado, mientras que el 20% manifiesta no contar con una mesa o escritorio para llevar acabo sus estudios.

**15. ¿Sus padres, fomentaban y estimulan sus estudios?**



**Tabla 52: Fomento de estudios-Estudiantes Desertores**

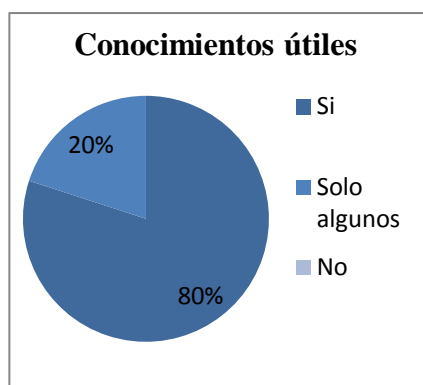
<b>Los padres fomentan y estimulan sus estudios</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	2	40%
No	3	60%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 60% de los estudiantes manifiesta no tener el apoyo de sus padres para sus estudios mientras que el 40% sostiene que sus padres fomentan y estimulan sus estudios.

### **16. ¿Considera que los conocimientos que adquirió durante su permanencia en la Carrera fueron de utilidad para la vida?**



**Tabla 53: Conocimientos adquiridos-Estudiantes Desertores**

<b>Conocimientos útiles para la vida</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	4	80%
Solo algunos	1	20%
No	0	0%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 80% de los estudiantes desertores considera que los conocimientos adquiridos durante su permanencia en la carrera son útiles para la vida, en tanto que el 20% sugiere que solo algunos conocimientos son relevantes.

### 17. ¿Fue alguna asignatura especialmente difícil para usted?

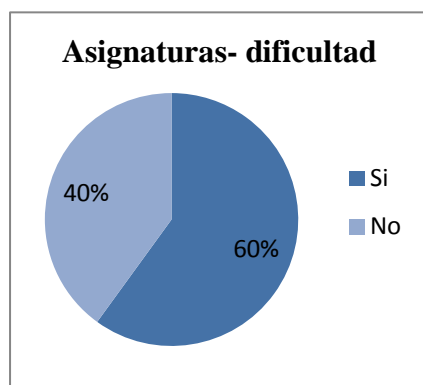


Tabla 54: Dificultad en las asignaturas-Estudiantes Desertores

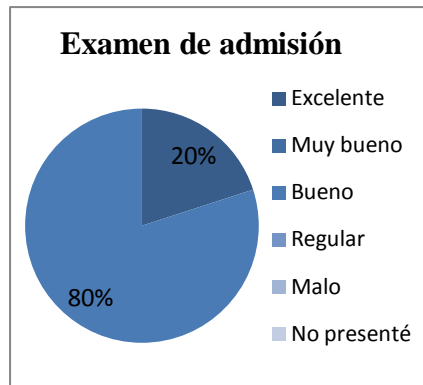
Asignatura que presentan dificultad	Frecuencia	Porcentaje
Si	3	60%
No	2	40%
<b>Cual:</b> Programación, Lógica matemática		
<b>Total</b>	5	100%

Elaborado por: Blanca Cuji

## Análisis e interpretación

Materias como Programación y Lógica Matemática son consideradas con tiene mayor dificultad por el 60% de los estudiantes desertores.

**18. En el examen de admisión de ingreso a la universidad que nota obtuvo:**



**Tabla 55: Examen de admisión**

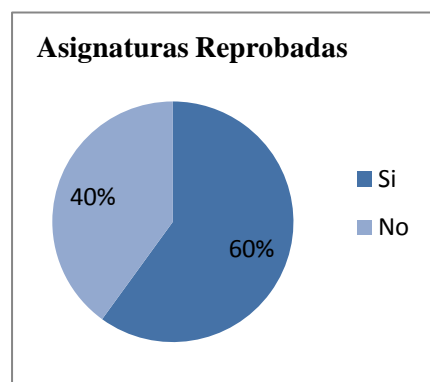
Examen de admisión	Frecuencia	Porcentaje
Excelente	1	20%
Muy bueno	0	0
Buena	4	80%
Regular	0	0
Malo	0	0
No presenté	0	0
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

La nota que obtuvieron en el examen de admisión a la universidad es buena según el 80% de estudiantes, y excelente para el otro 20%.

**19. ¿Durante su permanencia en la Carrera reprobó alguna asignatura de manera que haya retrasado algún periodo académico?**



**Tabla 56: Materias reprobadas-Estudiantes Desertores**

<b>Asignaturas reprobadas</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	3	60%
No	2	40%
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 60% de los estudiantes reprobaron por lo menos una materia durante su permanencia en la carrera.

### **20. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?**



**Tabla 57: Abandono de estudios-Estudiantes Desertores**

<b>Abandono de estudios</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	5	100%
No	0	0
<b>Total</b>	<b>5</b>	<b>100%</b>

Elaborado por: Blanca Cuji



**Si su respuesta fue positiva escoja cuál de estos factores influyeron en su decisión**

Económicos	5	100%
Rendimiento académico		
Sociales (Migración, religión, nivel económico)		
Demográfico (Edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)		
<b>Total</b>	<b>5</b>	<b>100%</b>

### **Análisis e interpretación**

El rendimiento académico y los factores económicos son las causas de la deserción de estudiantil en la totalidad de la población encuestada.

## **Análisis de la Encuesta realizada a docentes de la Carrera de Docencia en Informática**

### **I. DATOS GENERALES**

**Tabla 58: Edad-Docentes**

<b>Edad:</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
16-20	0	0
21-25	0	0
26-35	2	20%
Más de 36	8	80%
<b>Total</b>	<b>10</b>	<b>100</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

La edad de los docentes de la Carrera de Docencia en Informática está comprendida entre los 26 y más de 36 años.

## II CUESTIONARIO

### 1. ¿Cuál de las siguientes razones propician la Deserción Estudiantil?

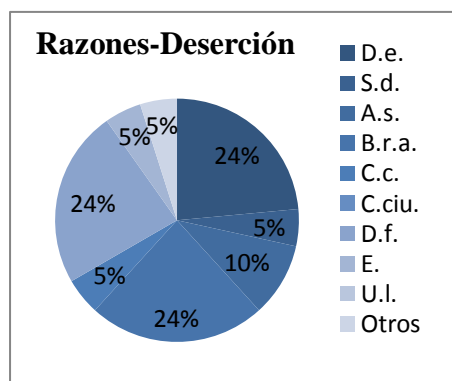


Tabla 59: Razones de Deserción Estudiantil-Docentes

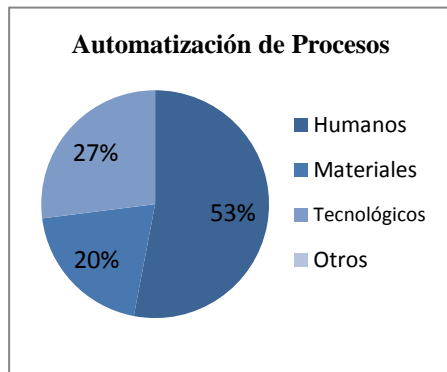
Razones que propician la Deserción Estudiantil	Frecuencia	Porcentaje
Dificultades económicas	5	24
Sanción disciplinaria	1	5
Aplazamiento de semestre	2	10
Bajo rendimiento académico	5	24
Cambio de carrera	1	5
Cambio de ciudad	0	0
Dificultades familiares	5	24
Enfermedad	1	5
Ubicación laboral	0	0
Otros	1	5
<b>Total</b>	<b>21</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El bajo rendimiento las dificultades económicas y familiares son a criterio de los docentes las razones de la Deserción Estudiantil.

**2. La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos:**



**Tabla 60: Automatización de procesos -Docentes**

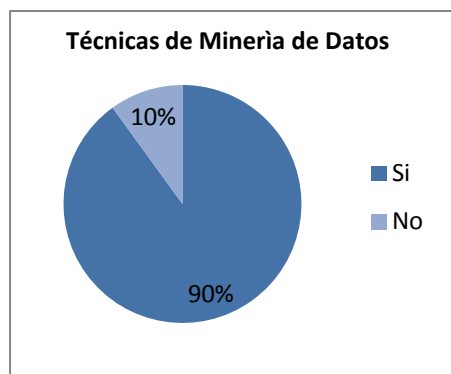
Estudios sobre Deserción Estudiantil	Frecuencia	Porcentaje
Humanos	8	53%
Materiales	3	20%
Tecnológicos	4	27%
Otros		0%
<b>Total</b>	<b>15</b>	<b>100%</b>

Elaborado por: Blanca Cuji

**Análisis e interpretación**

La automatización de procesos optimizará el uso de recursos humanos 53%, tecnológicos 20% y materiales 27% según el criterio de la gran mayoría de estudiantes.

**3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?**



**Tabla 61: Técnicas de minería de datos-Doctentes**

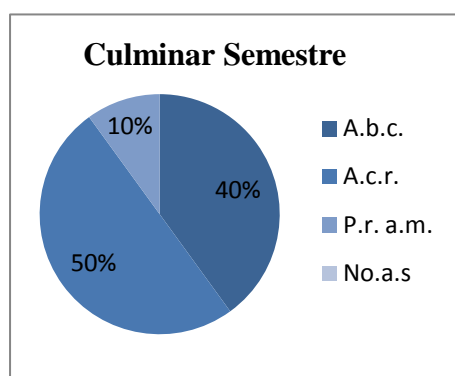
Técnicas de minería de datos	Frecuencia	Porcentaje
Si	9	90%
No	1	10%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 90% de los docentes considera que aplicar técnicas de minería de datos ayudará a detectar patrones de Deserción Estudiantil.

### **4. Con respecto a sus estudiantes, al culminar el semestre Octubre 2015 – Marzo 2016:**



**Tabla 62: Culminación de semestre-Doctentes**

Culminar el semestre	Frecuencia	Porcentaje
Aprobarán todas sus materias con buenas calificaciones.	4	40%
Aprobarán todas sus materias con calificaciones regulares	5	50%

Probablemente reprobren alguna materia	1	10%
No creo que aprueben este semestre	0	0
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

Los docentes consideran que al finalizar el periodo académico Octubre 2015 – Marzo 2016 el 50% de los estudiantes aprobarán sus materias con calificaciones regulares, mientras que el 40% lo harán con buenas calificaciones.

### 5. ¿Cuánto tiempo considera que los estudiantes, dedican a sus estudios como parte de su trabajo autónomo (fuera de clases)?

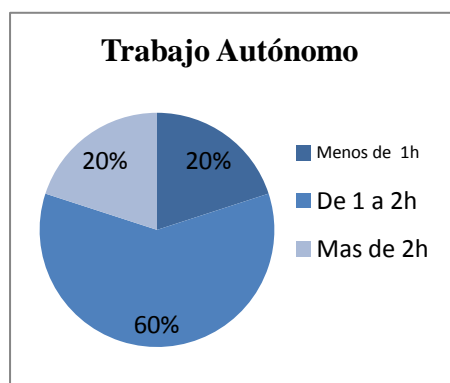


Tabla 63: Trabajo Autónomo-Docentes

Tiempo de dedicación - trabajo autónomo	Frecuencia	Porcentaje
Menos de 1h	2	20%
De 1 a 2h	6	60%
Más de 2h	2	20%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

De una a dos horas dedican los estudiantes a su trabajo autónomo así lo se señala el 60% de los docentes, el 20% señala que menos de una hora y otro 20% más de dos horas.

### 6. En cuanto al nivel económico de sus estudiantes es:

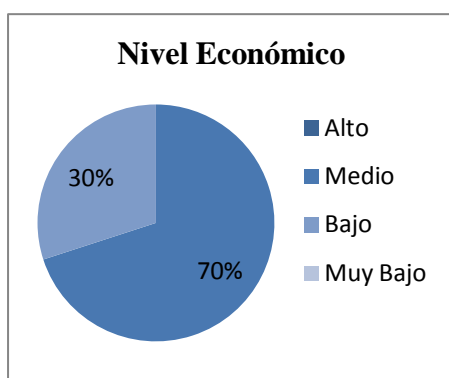


Tabla 64: Nivel Económico-Docentes

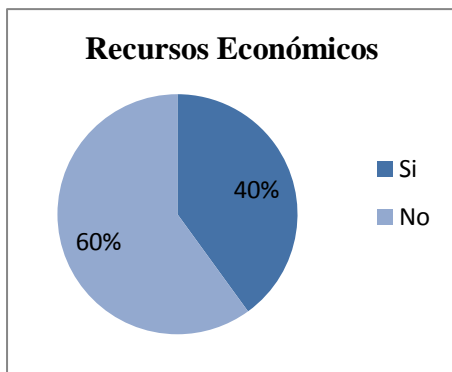
Nivel económico de estudiantes	Frecuencia	Porcentaje
Alto	0	0
Medio	7	70%
Bajo	3	30%
Muy Bajo	0	0
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

Los docentes consideran que el nivel económico de los estudiantes es de medio a bajo.

**7. ¿Considera que los estudiantes disponen de los recursos económicos para costear todos los gastos que se necesita para estudiar?**



**Tabla 65: Recursos Económicos**

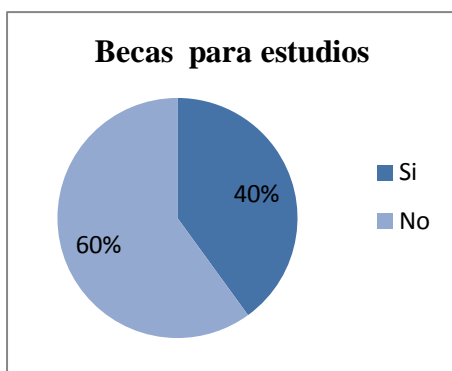
Recursos económicos	Frecuencia	Porcentaje
Si	6	60%
No	4	40%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### **Análisis e interpretación**

El 60% de los docentes manifiesta que los estudiantes si pueden costear los gastos de sus estudios, el 40% opina lo contrario.

**8. ¿Sus estudiantes cuentan con algún tipo de ayuda o beca para realizar sus estudios?**



**Tabla 66: Becas -Docentes**

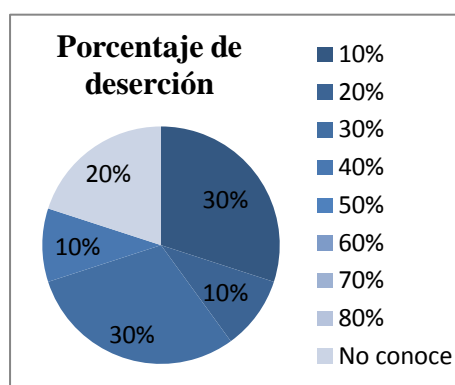
Beca para realizar los estudios	Frecuencia	Porcentaje
Si	4	40%
No	6	60%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### Análisis e interpretación

El 60% de los docentes considera que los estudiantes no reciben ningún tipo de ayuda o beca para sí estudios en tanto que el 40% considera que si reciben un beneficio

### 9. Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia?



**Tabla 67: Porcentaje de deserción-Docentes**

Porcentaje de deserción	Frecuencia	Porcentaje
10%	3	30%
20%	1	10%
30%	3	30%
40%	1	10%
50%	0	0
60%	0	0
70%	0	0
80%	0	0
No conoce	2	20%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji



## Análisis e interpretación

El 30% de los docentes considera que la deserción escolar se da 10% a 30% el 10% de los docentes la considera en un 40%; existe un 20% de docentes que desconocen sobre el porcentaje de Deserción Estudiantil.

### 10. Los estudiantes abandonan sus estudios cuando cursan el...de sus créditos.

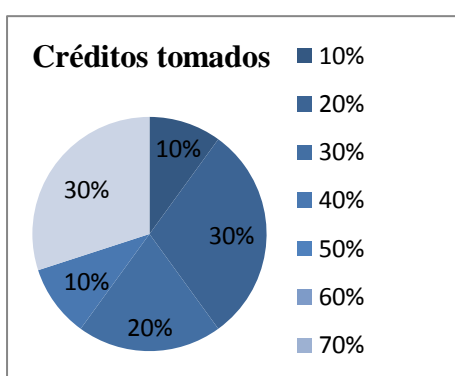


Tabla 68: Créditos tomados

Créditos tomados	Frecuencia	Porcentaje
10%	1	10%
20%	3	30%
30%	2	20%
40%	1	10%
50%	0	0
60%	0	0
70%	0	0
80%	0	0
No conoce	3	30%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

Los docentes consideran que los estudiantes abandonan sus estudios entre el 10% y 30% de sus créditos. Existe un elevado porcentaje de docentes que desconoce el número de créditos que los estudiantes toman al abandonar sus estudios

### 11. ¿Considera que los conocimientos que han adquirido los estudiantes en la institución son de utilidad para la vida?



Tabla 69: Conocimientos adquiridos-Docentes

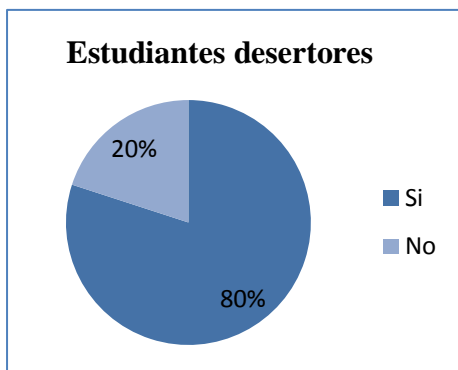
Conocimientos útiles	Frecuencia	Porcentaje
Si	5	50%
Solo Algunos	5	50%
No	0	0%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Análisis e interpretación

El 50% de los docentes manifiesta que el conocimiento que adquieren sus estudiantes es útil para la vida, el 50% restante considera que solo algunos conocimientos son de utilidad.

**12. ¿En su práctica docente ha tenido estudiantes que han desertado?**



**Tabla 70: Estudiantes Desertores-Docentes**

<b>Estudiantes desertores</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Si	8	80%
No	2	20%
<b>Total</b>	<b>10</b>	<b>100%</b>

Elaborado por: Blanca Cuji

Si su respuesta fue positiva escoja cuál de estos factores considera que influyeron en su decisión.

**Tabla 70.1: Estudiantes Desertores-Docentes**

Económicos y rendimiento académico	6	75%
Sociales (Migración, religión)	1	12,5%
Demográfico (Edad, género, estado civil, trabajo, lugar de residencia, etnia)	1	12,5%
<b>Total</b>	<b>8</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## **Análisis e interpretación**

El 80% de los docentes han tenido estudiantes que abandonaron sus estudios de ellos el 75% considera las razones de la deserción a factores económicos y rendimiento académico, en menor porcentaje situaciones como migración, religión edad, género, estado civil, lugar de residencia y etnia.

### **4.2. Verificación de la Hipótesis**

#### **4.2.1. Planteamiento de la Hipótesis**

Las técnicas de predicción incidirán en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato.

- **Modelo Lógico**

**H0:**Las técnicas de predicción NO inciden en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato.

**H1:**Las técnicas de predicción SI inciden en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato.

- **Modelo Estadístico:**

Se utilizó la prueba estadística Chi - Cuadrado, para determinar si existe o no, la relación entre las dos variables. Para lo cual se utilizó la siguiente fórmula:

$$\chi^2 = \sum \frac{(\chi - \chi)^2}{\chi}$$

Donde:

$\chi^2$  = Chi – Cuadrado

O = Frecuencias Observadas

E = Frecuencias Esperadas

**Nivel de Significancia**  $\alpha = 0,05$

**Nivel de confianza** = 95 %

Para la verificación de la hipótesis, se eligió dos preguntas por cada variable para determinar la relación que existe entre: Las Técnicas de Predicción y la Deserción Estudiantil, las preguntas y respuestas seleccionadas fueron:

**Variable: Técnicas de Predicción**

3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?

**Tabla 71.1: Técnicas de minería de datos-Docentes**

Opciones	Frecuencia	Porcentaje
Si	59	84%
No	11	16%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

17. ¿Es alguna asignatura especialmente difícil para Usted?

**Tabla 26.1: Dificultad de las asignaturas**

Opciones	Frecuencia	Porcentaje
Si	28	40%
No	42	60%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

## Variable: Deserción Estudiantil

19. ¿Ha reprobado alguna asignatura de manera que se haya retrasado en algún periodo académico?

**Tabla 28.1: Reprobación de asignaturas**

Opciones	Frecuencia	Porcentaje
Si	19	27%
No	51	73%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

20. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?

**Tabla 29.1: Abandono de estudios**

Opciones	Frecuencia	Porcentaje
Si	25	36%
No	45	64%
<b>Total</b>	<b>70</b>	<b>100%</b>

Elaborado por: Blanca Cuji

### 4.2.2. Tabla de Frecuencias Observadas

#### 4.2.2: Resumen de preguntas

Variables	Preguntas	Frecuencias Observadas		Total
		Si	No	
VI	3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?	59	11	70
	17. ¿Es alguna asignatura especialmente difícil para Usted?	28	42	70

VD	19. ¿Ha reprobado alguna asignatura de manera que se haya retrasado en algún periodo académico?	19	51	70
	20. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?	25	45	70
<b>Total</b>		131	149	<b>280</b>

Elaborado por: Blanca Cuji

#### 4.2.3. Tabla de Frecuencias Esperadas

Tabla 4.2.3: Frecuencias esperada

Variables	Preguntas	Frecuencias Observadas		Total
		Si	No	
VI	3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?	32,75	37,25	70
	17. ¿Es alguna asignatura especialmente difícil para Usted?	32,75	37,25	70
VD	19. ¿Ha reprobado alguna asignatura de manera que se haya retrasado en algún periodo académico?	32,75	37,25	70
	20. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?	32,75	37,25	70
<b>Total</b>		131	149	280

Elaborado por: Blanca Cuji

#### 4.2.4. Prueba Chi Cuadrado

**Tabla 4.2.4: Técnicas de Predicción**

Variable	Alternativas	O	E	(O-E)	(O-E) <sup>2</sup>	(O-E) <sup>2</sup> /E
VI	Si	59	32,75	26,25	689,06	21,04
	Si	28	32,75	-4,75	22,56	0,69
	No	11	37,25	-26,25	689,06	18,50
VD	No	42	37,25	4,75	22,56	0,61
	Si	19	32,75	-13,75	189,06	5,77
	Si	25	32,75	-7,75	60,06	1,83
	No	51	37,25	13,75	189,06	5,08
	No	45	37,25	7,75	60,06	1,61
					$\chi^2 =$	<b>55,13</b>

Elaborado por: Blanca Cuji

El Chi - Cuadrado Calculado es **55,13**.

#### 4.2.5. Grados de Libertad

$$gl = (f - 1) (c - 1)$$

$$gl = (4 - 1) (2 - 1)$$

$$gl = 3 * 1$$

$$gl = 3$$

Con tres grados de libertad y 95% de nivel de confianza el valor Chi – Cuadrado teórico o tabular según la tabla de distribución es: **7,8147** según se muestra en la siguiente Figura.

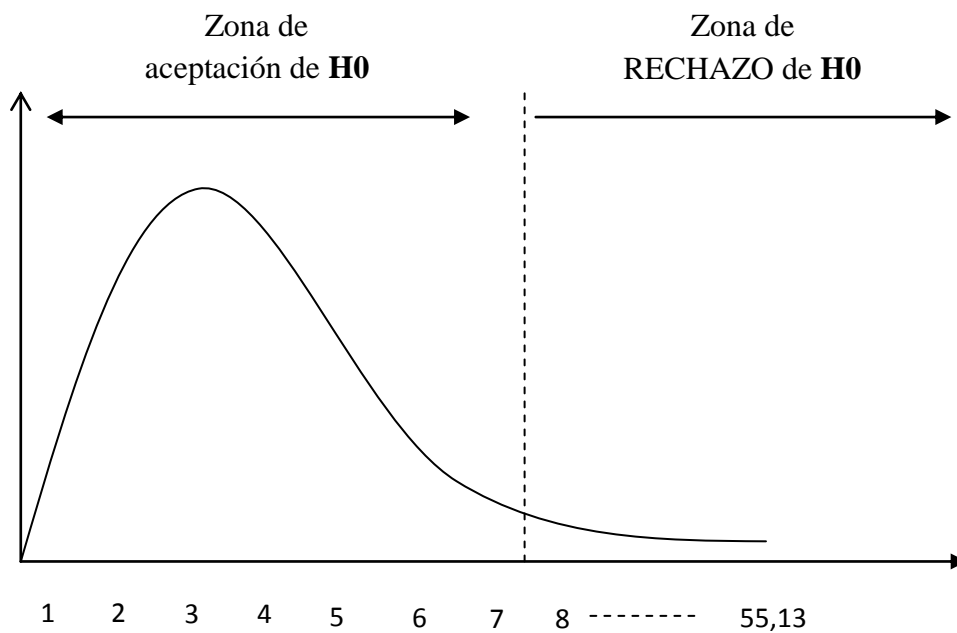
v/p	0,001	0,0025	0,005	0,01	0,025	0,05
1	10,8274	9,1404	7,8794	6,6349	5,0239	3,8415
2	13,8150	11,9827	10,5965	9,2104	7,3778	5,9915
3	16,2660	14,3202	12,8381	11,3449	9,3484	<b>7,8147</b>

Valor Chi - Cuadrado según la tabla de distribución

**Fuente:** [http://labrad.fisica.edu.uy/docs/tabla\\_chi\\_cuadrado.pdf](http://labrad.fisica.edu.uy/docs/tabla_chi_cuadrado.pdf) 56



## Regla de decisión



Si  $X^2$  calculado  $>$   $X^2$  tabular entonces se rechaza la hipótesis nula ( $H_0$ ) y se acepta la hipótesis alterna  $H_1$ .

Por tanto:

$$X^2 \text{ calculado} > X^2 \text{ tabular}$$

$$55,13 > 7,8147$$

Debido a que el Chi cuadrado Calculado es mayor que el Chi cuadrado tabular se comprueba la hipótesis Alterna que dice **H1**: Las técnicas de predicción incidirán en la detección de patrones de Deserción Estudiantil en la Carrera de Docencia en Informática de la Facultad de Ciencias Humanas y de la Educación de la Universidad Técnica de Ambato, y se rechaza la hipótesis nula ( $H_0$ ).

## CAPÍTULO V

### CONCLUSIONES Y RECOMENDACIONES

#### 5.1. Conclusiones

- Las dificultades económicas y familiares, así como el bajo rendimiento son factores predominantes en la Deserción Estudiantil, poca incidencia tiene indicadores como: sanción disciplinaria, cambio de ciudad, ubicación laboral, cambio de carrera, espacio físico para los estudios, nota obtenida en el examen de admisión, influencia de los padres al momento de elegir la Carrera y enfermedad.
- Aproximadamente el 24,3% de los estudiantes tienen un nivel económico entre bajo y muy bajo, solventan sus estudios con recursos propios, conforme a este indicador probablemente 17 de los 70 estudiantes que cursan la carrera de Docencia en Informática serían posibles desertores.
- Las escasas horas que los estudiantes dedican al trabajo autónomo, así como créditos retrasados, semestres reprobados ,poca utilidad de los conocimientos que adquieren en el aula, materias consideradas “difíciles” como Programación, Base de Datos y Física, Lógica Matemática afecta al rendimiento académico de los estudiantes y elevan los riesgos de deserción.
- Existe un porcentaje elevado de desconocimiento por parte de la población encuestada sobre el porcentaje de Deserción Estudiantil que se da en la Carrera, así como del número de créditos que logran alcanzar los estudiantes antes de abandonar sus estudios.
- La deserción de los estudiantes es el resultado de la combinación
- y efectos de distintas variables, las cuales muchas veces son independientes y no se pueden controlar como el nivel económico.

## 5.2. Recomendaciones

- Identificados los indicadores de deserción, como número de horas de trabajo autónomo, dificultad en diferentes asignaturas es necesario incrementar las horas de trabajo autónomo y tutoría académica con énfasis en materias como Programación, Bases de Datos y Física.
- Automatización de procesos en los estudios sobre Deserción Estudiantil para optimizar recursos humanos, materiales y tecnológicos.
- Diseñar un Modelo de predicción basado en técnicas de minería de datos(clasificación), para predecir una posible Deserción Estudiantil.
- Socialización de las tasas de Deserción Estudiantil entre docentes, estudiantes y autoridades de la Facultad.

## CAPÍTULO VI

### PROPUESTA

#### 6.1. Tema:

Modelo de Predicción de Deserción Estudiantil basado en Técnicas de Clasificación usando Arboles de Decisión para la Carrera de Docencia en Informática

#### 6.2. Datos Informativos

**Institución Ejecutora:** UTA – Facultad de Ciencias Humanas y de la Educación  
Carrera de Docencia en Informática.

**Dirección:** Av. Los Chasquis y Guayllabamba

**Beneficiarios:** Estudiantes de la Carrera de Docencia en Informática.

**Tiempo:** 2015 - 2016

**Costo:** El costo estimado para el desarrollo y ejecución de la propuesta es 3000USD

**Tutor:** Ing. Mg. Edison Álvarez

#### 6.3. Antecedentes de la propuesta

Se toma como antecedente los modelos realizados en diferentes investigaciones, las cuales se mencionan a continuación: “*La minería de datos como un método innovador para la detección de patrones de deserción estudiantil en programas de pregrado en instituciones de educación superior*” (Timarán & Calderón, 2013); en esta investigación se detectó patrones de Deserción Estudiantil a partir de datos socioeconómicos, académicos, disciplinares e institucionales, utiliza la

metodología de las etapas del proceso de descubrimiento de conocimiento en bases de datos, a partir del año 2014 dándole un seguimiento hasta el año 2011, determinando si desertaron o no, se creó un repositorio de datos para la limpieza de la información y se aplicó la técnica de clasificación basada en árboles de decisión usando R. Los resultados fueron analizados, evaluados e interpretados para determinar la validez del conocimiento obtenido.

*“Factores determinantes que inciden en la deserción de los estudiantes universitarios”* (Apaza, 2012) aborda el tema de la deserción de los estudiantes en la educación Superior Universitaria basado en aspectos psicológicos, económicos, sociológicos, organizacionales, mencionados a dichos como predictores para el abandono estudiantil, poniendo énfasis en variables más críticas.

*“Improving student success using predictive models and data visualisations”* (Essax & Ayad, 2012) el autor se refiere al porcentaje de estudiantes desertores en Estados Unidos en el nivel universitario aproximadamente el 50% de la totalidad de estudiantes que ingresa a la universidad logra graduarse de 4 a 5 años, para esto se propuso un modelo predictivo que utiliza la inteligencia artificial y estadística para identificar estudiantes con riesgo de deserción.

*“La minería de datos compete con el corredor de bolsa”* (Hope, 2016) en donde se menciona un sin número de modelos predictores elaborados para comprar, vender o mantener acciones mediante el desarrollo de diversos algoritmos.

#### **6.4. Justificación**

Un indicador importante para medir la calidad de la educación es la tasa de retención que tienen las instituciones educativas, en tal virtud es necesario disminuir los índices de Deserción Estudiantil, la propuesta plantea la posibilidad de crear un modelo predictivo basado en datos socioeconómicos y académicos recolectados de la Carrera de Docencia en Informática a partir del año 2006 hasta

el año 2015, mediante el uso de técnicas de minería de datos previo un análisis de las variables que tienen mayor incidencia en la deserción de la Carrera.

## **6.5. Objetivos de la Propuesta**

### **Objetivo General**

- Diseñar un Modelo de Predicción de Deserción Estudiantil usando Árboles de Decisión.

### **Objetivos Específicos**

- Recolectar datos académicos y socioeconómicos de la Carrera de Docencia en Informática a partir del año 2006-2015.
- Limpiar los datos.
- Analizar las variables.
- Aplicar un método de clasificación usando árboles de decisión para la generación del modelo predictivo.
- Visualizarlos datos.

## **6.6. Análisis de Factibilidad**

La factibilidad en cuanto a la base legal se sustenta en el Plan Nacional del Buen Vivir enmarcada dentro del Objetivo 9: “Garantizar el trabajo digno en todas sus formas” inciso 9.2 literal e que manifiesta: “Priorizar el desarrollo de iniciativas económicas vinculadas al desarrollo de las TIC, aprovechando las capacidades desarrolladas en software y de los recursos de la biodiversidad, creando espacios e infraestructura pertinente, que sustenten su productividad”. (Consejo Nacional de Planificación, 2013).

Por otra parte, se cuenta con los recursos económicos necesarios por parte del investigador y la apertura de las autoridades de la Carrera de Docencia en Informática para la obtención de los datos.

Finalmente se cuenta con los conocimientos técnicos sobre estadística, bases de datos, matemática aplicada y técnicas de visualización, requeridos para la aplicación de la Minería de datos por parte del investigador.

## **6.7. Fundamentación**

KDD (Knowledge Discovery in Databases) nombre técnico con que se denomina al proceso global de extracción de conocimiento de bases de datos. (Villena & Crespo, 2012)

### **Etapas a llevar a cabo durante el proceso de Minería de Datos**

#### **Etapa de Selección de datos**

El objetivo de esta etapa es obtener las fuentes internas y externas de datos que sirven de base para el proceso de minería de datos. (Timarán & Calderón, 2013). El preprocesamiento de datos abarca procesos de extracción, limpieza e integración de la información.

#### **Etapa de procesamiento de datos**

Esta etapa es de larga duración y consume gran cantidad de recursos, por ende, requiere de mayor tiempo en el desarrollo de un proyecto de minería de datos. Aproximadamente el 70% o más de la duración estimada de un proyecto, se gasta en el análisis de datos. (Medina, 2011).

El objetivo de esta etapa es obtener datos limpios, sin valores nulos o anómalos que permitan obtener patrones de calidad. (Timarán & Calderón, 2013).

#### **Etapa de Transformación de datos**

El objetivo de esta fase es transformar la fuente de datos en un conjunto listo para aplicar las técnicas de minería de datos.

#### **Etapa de Minería de Datos**

El objetivo de esta etapa es la búsqueda y descubrimiento de patrones de interés aplicando técnicas de minería de datos para el descubrimiento de nuevo conocimiento, tales como clasificación, etc.

### **Etapas de Interpretación/ Evaluación de datos**

Etapas en las que se interpretan y evalúan los patrones descubiertos con el propósito de consolidar el conocimiento descubierto para su posterior incorporación en otro sistema para confrontarlo con conocimiento descubierto con anterioridad. (Timarán & Calderón, 2013).

### **Métodos de clasificación**

Análisis de conglomerados (cluster), cuyo objetivo es descubrir conjuntos de casos (clientes, ventas, productos...) o de variables que son similares y que se agrupan o tienen características similares. Esta segmentación define tipologías o clases de elementos "parecidos". Así, si dos elementos (casos o variables) se definen mediante la observación de  $p$  características, es necesario definir una medida de distancia o similitud entre ellos. Por ejemplo, la distancia euclídea. (Caridad, 2011)

$$d(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

Es posible establecer una clasificación mediante la definición de un conjunto de reglas independientes. No precisa establecer una división jerárquica con un esquema en árbol, aunque las reglas pueden dar lugar a situaciones contradictorias. En el proceso de clasificación es usual atribuir un nivel de confianza a cada regla. (Caridad, 2011)

### **R**

Es un entorno de software gratuito para cálculos y gráficos estadísticos proporciona un entorno sofisticado para minería de datos, análisis estadísticos y visualización de datos.

### **Rattle**

Interfaz gráfica de usuario para minería de datos, se usa bajo la Licencia pública general.



## **D3js**

Librería de JavaScript para producir, a partir de datos, infogramas dinámicos e interactivos en navegadores web. Hace uso de tecnologías bien sustentadas como SVG (Scalable Vector Graphics), HyperText Markup Language (Html5), y Cascading Style Sheets(CSS).

## **Conocimientos requeridos de otras ciencias para hacer minería de datos**

La minería de datos es la confluencia de muchas disciplinas por ello se requiere conocimientos de:

1. Estadística
2. Tecnología de bases de datos
3. Matemática Aplicada
4. Ciencias de la Información
5. Técnicas de Visualización

## **Método Predictivo o de aprendizaje supervisado**

Son supervisados porque se trata de entrenar modelos; mediante una supervisión, se crean modelos para detectar aspectos específicos como: fraudes, detectar si un cliente va ser buen pagador o no, si un estudiante desertará o no de sus estudios.

Para los modelos predictivos se usan tanto variables cualitativas como cuantitativas

## **6.8. Metodología**

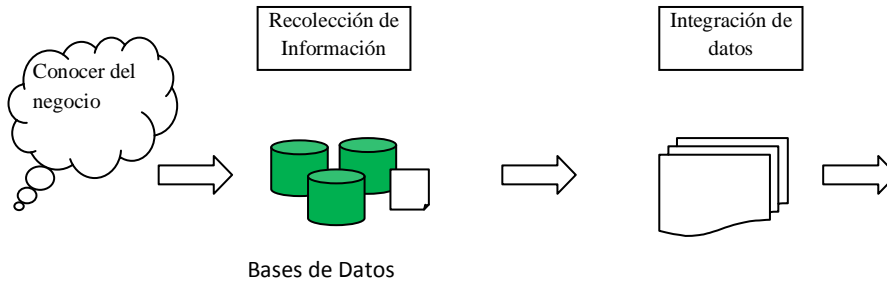
### **DESARROLLO DE LA PROPUESTA**

#### **Modelo de Predicción de Deserción Estudiantil basado en Técnicas de Clasificación usando Árboles de Decisión para la Carrera de Docencia en Informática**

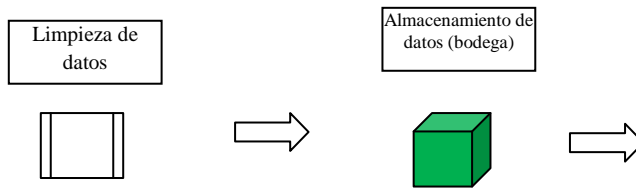
**Procedimiento a seguir:**

**Aplicación del Proceso KDD (Proceso de extracción del conocimiento)  
basado en las siguientes etapas:**

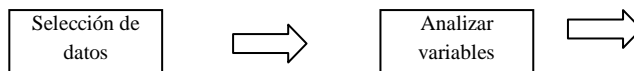
**1. Etapa de Selección de Datos**



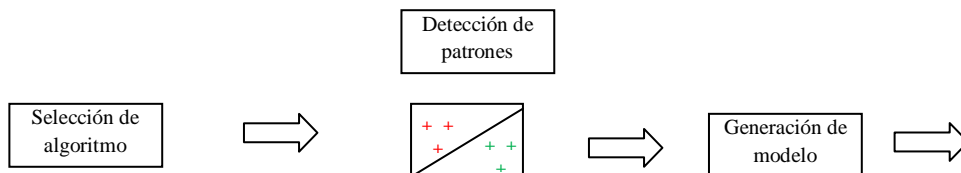
**2. Etapa de procesamiento de datos**



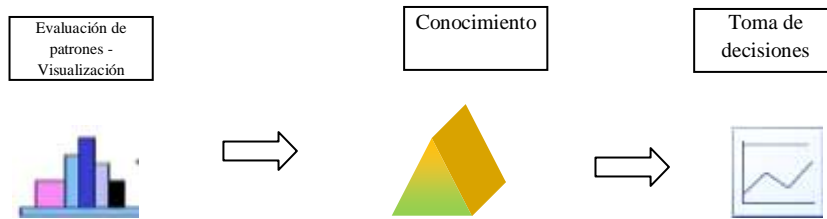
**3. Etapa de Transformación de datos**



**4. Etapa de Minería de Datos**



## 5. Etapa de Interpretación/ Evaluación de patrones



### ETAPA I

### SELECCIÓN DE DATOS

#### 1.1.Recolección de la Información

Los datos crudos corresponden a los estudiantes de la Carrera de Docencia en Informática recolectados a partir del año 2006y distribuidos en dos matrices de la siguiente manera:

**Matriz 1:** Fila de registros con diferentes datos y columnas con los siguientes atributos (Hoja 1 – documento de Excel):

1. Identificador
2. Apellido 1
3. Apellido 2
4. Nombres
5. Genero
6. Estado Civil
7. Etnia
8. Fecha de Nacimiento
9. Lugar nacimiento
10. Ciudad residencia

**Figura 1: Datos personales**

Identificador	Apellido	Nombres	Género	Estado Civil	Etnia	Fecha Nacimiento	Ciudad Nacimiento	Ciudad Residencia
1	A	GUERRERO	VICTOR XA	MASCULINO	SOLTERO	MESTIZA	08/10/1994	PALANDA
2	B	CARDENAS	WASHINGTON	MASCULINO	SOLTERO	MESTIZA	29/05/1996	SANTA CRUZ
3	C	PALATE	SERGIO LEI	MASCULINO	SOLTERO	MESTIZA	10/11/1988	PELLICO
4	D	REAL	GLADYS M	FEMENINO	SOLTERO	MESTIZA	30/10/1989	AMBATO
5	E	MAMALLACTA	MARICELA	FEMENINO	SOLTERO	MESTIZA	04/09/1988	MOCHA
6	F	LOPO	GUAMANAF	FEMENINO	SOLTERO	MESTIZA	26/11/1993	AMBATO
7	G	CHICAIZA	VALERIA A	FEMENINO	CASADO	MESTIZA	17/02/1992	PILLARO
8	H	AMARA	MARIA PA	FEMENINO	SOLTERO	MESTIZA	19/12/1988	PILLARO
9	I	ORTEGA	KATHERINE	FEMENINO	SOLTERO	MESTIZA	24/12/1995	PILLARO
10	J	CALL	DORA LILY	FEMENINO	SOLTERO	MESTIZA	20/12/1991	PILLARO
11	K	SAMBININO	SAMUEL B	MASCULINO	SOLTERO	MESTIZA	16/10/1995	PILLARO
12	L	AMACAÑA	SILVIA ALE	FEMENINO	SOLTERO	MESTIZA	28/01/1988	PILLARO
13	M	COLLAQUIZO	LUIS ISRAE	MASCULINO	CASADO	MESTIZA	15/07/1985	PILLARO
14	N	YANCHATIPANI	LILIANA CA	FEMENINO	SOLTERO	MESTIZA	09/12/1989	PILLARO
15	O	RODRIGUEZ	ERIKA MISH	FEMENINO	SOLTERO	MESTIZA	02/03/1995	BAÑOS
16	P	MASAGUIZA	JOSÉ GUST	MASCULINO	SOLTERO	INDIGENA	16/07/1990	PELLICO
17	Q	MASAGUIZA	SANTIAGO	MASCULINO	CASADO	INDIGENA	01/02/1988	PELLICO
18	R	VILLENA	HENRRY JA	MASCULINO	SOLTERO	MESTIZA	01/13/1989	PELLICO
19	S	CUYANGUILLO	ERIKA JESS	FEMENINO	SOLTERO	MESTIZA	16/06/1992	PELLICO
20	T	PAROLES	DIEGO SAN	MASCULINO	CASADO	MESTIZA	12/06/1988	PELLICO

Fuente: Datos Históricos a partir del año 2016

**Matriz 2:** Fila de registros con diferentes datos y columnas con los siguientes atributos ((Hoja 2 – documento de Excel):

1. Identificador
2. Periodo Académico
3. Curso Matriculado
4. Materia
5. Nota 1
6. Nota 2

**Figura 2 : Notas**

Identificador	Periodo	Curso Matriculado	Materia	Nota 1	Nota 2
1	SEP/05-FEB/06	QUINTO SEMESTRE A	CÁLCULO INTEGRAL	7	9
2	SEP/05-FEB/06	QUINTO SEMESTRE A	DIDÁCTICA APLICADA	7,6	7,7
3	SEP/05-FEB/06	QUINTO SEMESTRE A	EVALUACIÓN DE LOS APRENDIZAJES	8	9
4	SEP/05-FEB/06	QUINTO SEMESTRE A	LENGUAJE VISUAL I	5,6	7,5
5	SEP/05-FEB/06	QUINTO SEMESTRE A	BASE DE DATOS I	9,5	6,1
6	SEP/05-FEB/06	QUINTO SEMESTRE A	INFORMÁTICA APLICADA II	7	8
7	SEP/05-FEB/06	QUINTO SEMESTRE A	CÁLCULO INTEGRAL	8	7
8	SEP/05-FEB/06	QUINTO SEMESTRE A	DIDÁCTICA APLICADA	7,5	8,7
9	SEP/05-FEB/06	QUINTO SEMESTRE A	EVALUACIÓN DE LOS APRENDIZAJES	9	9
10	SEP/05-FEB/06	QUINTO SEMESTRE A	LENGUAJE VISUAL I	6,3	8,8
11	SEP/05-FEB/06	QUINTO SEMESTRE A	BASE DE DATOS I	7,3	9,2
12	SEP/05-FEB/06	QUINTO SEMESTRE A	INFORMÁTICA APLICADA II	9	8
13	SEP/05-FEB/06	QUINTO SEMESTRE A	CÁLCULO INTEGRAL	9	9
14	SEP/05-FEB/06	PRIMER SEMESTRE A	UTILITARIOS	8	8,4
15	SEP/05-FEB/06	PRIMER SEMESTRE A	INFORMÁTICA	8,5	7,9
16	SEP/05-FEB/06	PRIMER SEMESTRE A	MATEMÁTICAS BÁSICAS	7,5	7
17	SEP/05-FEB/06	PRIMER SEMESTRE A	TÉCNICAS DE INVESTIGACIÓN	8	10
18	SEP/05-FEB/06	PRIMER SEMESTRE A	LENGUAJE Y COMUNICACIÓN	9,1	10

Fuente: Datos Históricos a partir del año 2016

## 1.2.Integración de datos

Originalmente el número de registro con los que se contó para el análisis fue de 484, se integró en una única matriz los atributos de la Matriz 1+ Matriz 2.

Figura 3: Integración de datos

The screenshot shows an Excel spreadsheet with the following columns: APELLIDO, NOMBRE, SEXO, ESTADO CIVIL, ETNIA, NACIMIENTO, LUGAR DE NACIMIENTO, EDAD, FECHAS DE INGRESO, LUGAR DE INGRESO, CICLO DE INGRESO, NIVEL, CURSOS, and a series of columns for academic notes (NOTA1P, NOTA1S, etc.). The data rows contain student records with their respective details and scores.

Fuente: Datos personales + Notas

Se agregó la columna *Edad*, *Desertor*, *Nota1p*, *Nota2p*, *Nota2s*, *Nota2s*

- *Edad*: Edad de ingreso a la carrera; se determinó con la columna fecha de nacimiento y fecha de ingreso.
- *Desertor*: Se determinó mediante la columna *Curso Matriculado* conjuntamente con el número de notas por materia que conste en el histórico de datos del estudiante.
- *Nota1p*, *Nota2p*, *Nota2s*, *Nota2s*: Promedio de notas de todas las materias tomadas por los estudiantes en el primer semestre (*Nota1p*, *Nota2p*) y segundo semestre *Nota2s*, *Nota2s*.

Se tomaron las notas del primero y segundo semestre por ser los niveles en los cuales existe mayor cantidad de estudiantes desertores; son los niveles en los cuales se evidencia el mayor porcentaje de deserción.

**Tabla 72: Porcentaje – Deserción**

DESERTORES a partir del periodo sep05-fer/06 hasta oct./15-feb/2016	Cantidad	Porcentaje
SI	169	35%
NO	315	65%
<b>Total</b>	<b>484</b>	<b>100%</b>

**Fuente:** Datos Históricos a partir del año 2016

**Tabla 73 : Número de desertores por semestre**

Semestre	Cantidad
Primero	68
Segundo	48
Tercero	19
Cuarto	13
Quinto	4
Sexto	6
Séptimo	2
Octavo	1
No hay datos	8
<b>Total</b>	<b>169</b>

**Fuente:** Datos Históricos a partir del año 2016

## ETAPA II

### PROCESAMIENTO DE DATOS

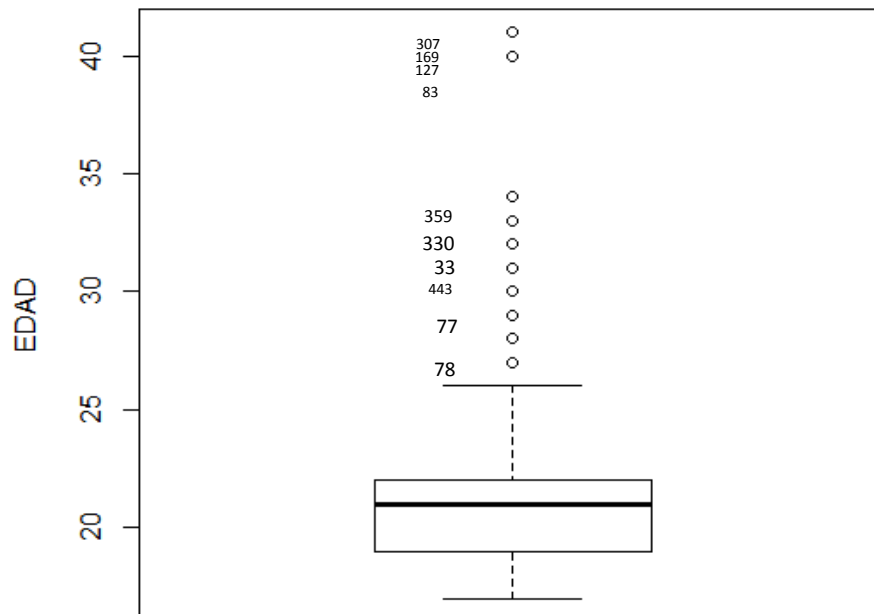
#### 2.1. Limpieza de datos

Para la limpieza de datos se partió con 484 estudiantes de la Carrera de Docencia en informática a partir del periodo Sep/2005 – Feb/2006. Se detectó **DATOS ATÍPICOS** utilizando el programa FactorMiner de R; mediante diagramas de pareto para las variables *Edad*, *Nota1p*, *Nota2p*, *Nota2s*, *Nota2s*.

- **Datos Atípicos variable *Edad***

Boxplot( ~ EDAD, data=Datos, id.method="y")

[1] "127" "169" "307" "83" "359" "330" "33" "443" "77" "78"

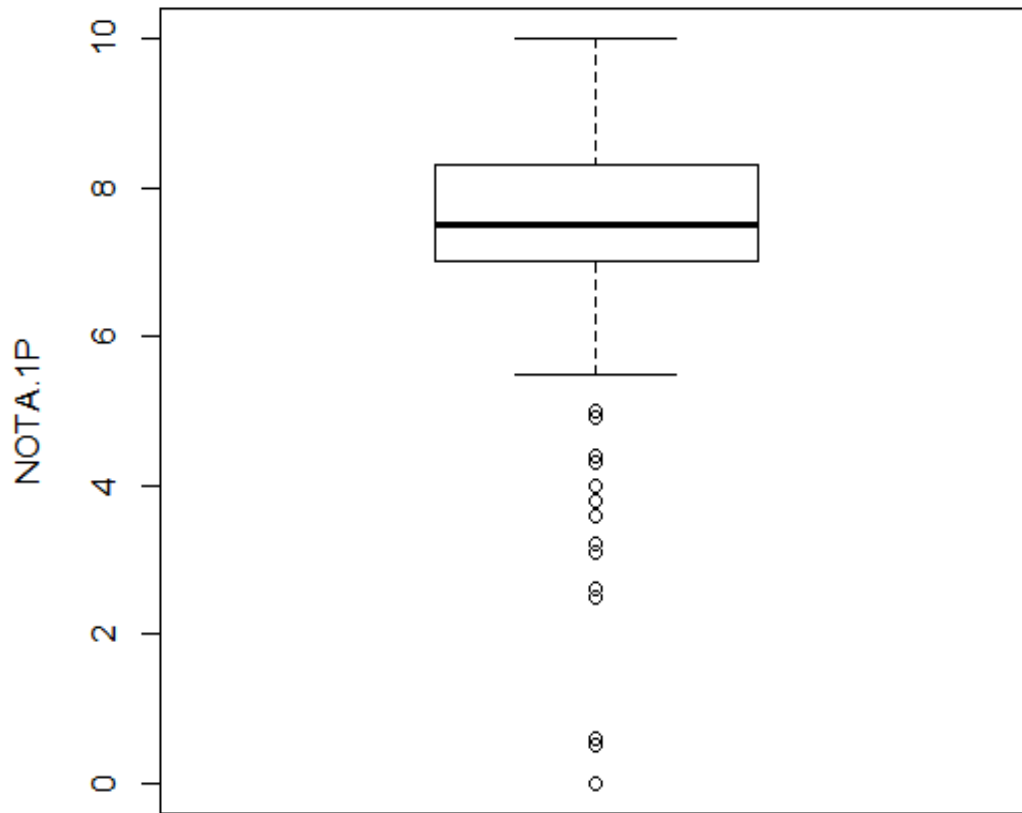


Tres estudiantes tienen 41 años; un estudiante 40; un estudiante 34 años; un estudiante 33 años; dos estudiantes 32 años; un estudiante con 31 años y finalmente un estudiante con 30 años.

- **Datos Atípicos variable *Nota1p***

Boxplot( ~ NOTA.1P, data=Datos, id.method="y")

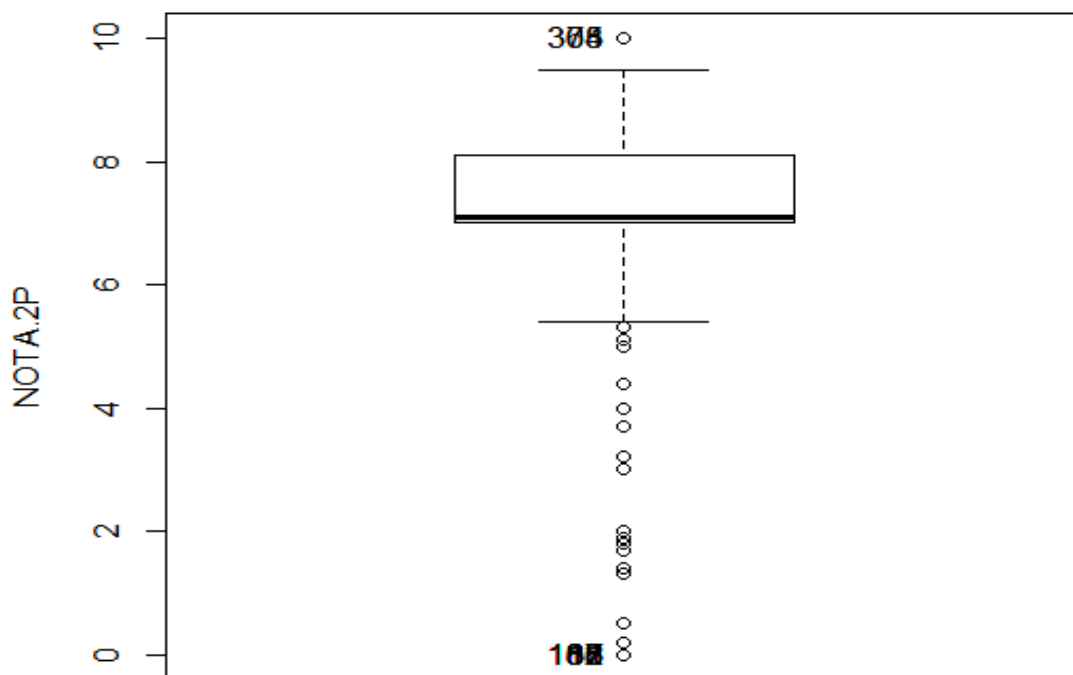
[1] "11" "82" "88" "168" "185" "187" "212" "221" "235" "252"



- **Datos Atípicos variable *Nota2p***

Boxplot( ~ NOTA.2P, data=Datos, id.method="y")

[1] "11" "16" "82" "87" "88" "102" "133" "168" "185" "187" "78"  
"305" "364"

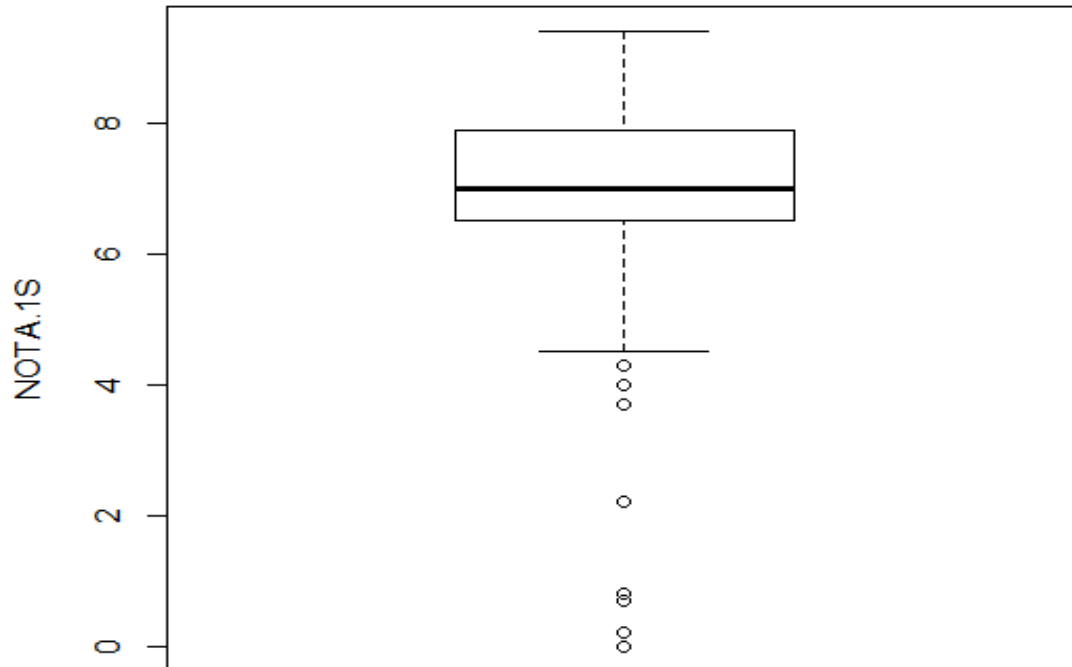




**Datos Atípicos variable *Nota1s***

Boxplot( ~ NOTA.1S, data=Datos, id.method="y")

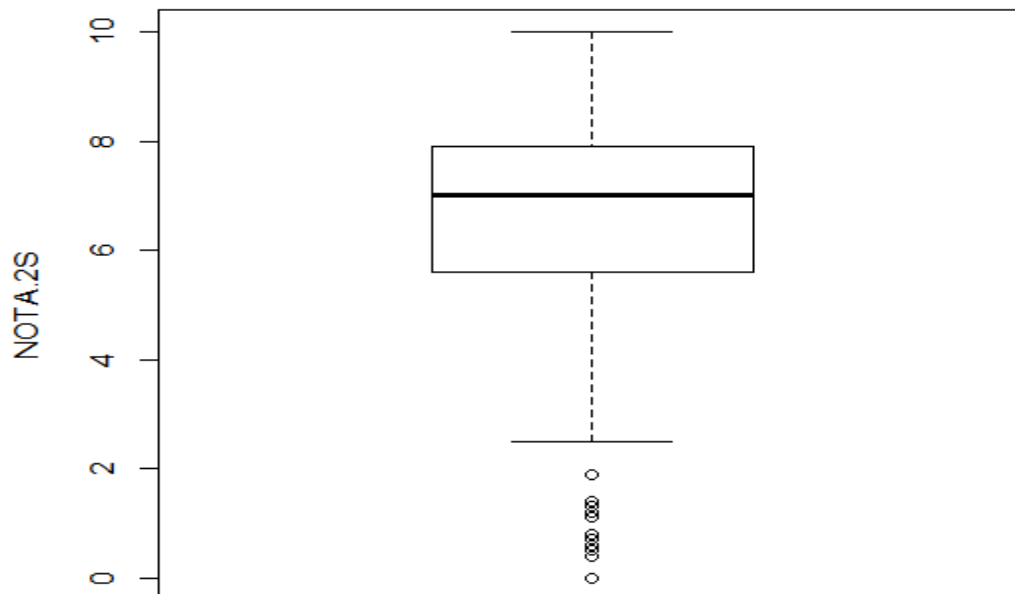
[1] "7" "11" "19" "37" "40" "77" "82" "87" "88" "102"



• **Datos Atípicos variable *Nota2s***

Boxplot( ~ NOTA.2S, data=Datos, id.method="y")

[1] "7" "9" "11" "19" "29" "37" "40" "66" "77" "82"



- Posterior a detectar los datos atípicos, se retiró de la base de datos, a 24 estudiantes que mostraron atipicidades en las variables *Nota1p*, *Nota2p*, *Nota2s*, *Nota2s*. pues muestran notas de 0,0 en las cuatro variables
- Se reemplazó la palabra *Baños* por *Banios* de las variables lugar de nacimiento y ciudad de residencia.
- Además, se retiró a ocho estudiantes por poseer solo matricula en determinado periodo académico y no contar con notas en ningún semestre

IDENTIFICADOR	APELLIDO 1	GENERO	ESTADO CIVIL	ETNIA	FECHA NACIMIENTO	FECHA INGRESO A LA CARRERA	EDAD	PERIODO DE INGRESO	LUGAR NACIMIENTO	CIUDAD RESIDENCIA	DESCRIPCIÓN	NOTA P1	NOTA P2
100	ARDILA	MASCULINO	SOLTERO	MESTIZA	04/07/1988			-90	No existe dato	LATAJUNGA	AMBATO	SI	NO HAY DATOS NO HAY DATOS
199	GUAMBO	MASCULINO	SOLTERO	MESTIZA	10/02/1977			-76	No existe dato	PELLEO	PELLEO	SI	NO HAY DATOS NO HAY DATOS
280	TORRES	FEMENINO	SOLTERO	MESTIZA	04/07/1988			-90	No existe dato	AMBATO	AMBATO	SI	NO HAY DATOS NO HAY DATOS
309	MARTINEZ	MASCULINO	SOLTERO	MESTIZA	26/08/1983			-85	No existe dato	PULLU	PULLU	SI	NO HAY DATOS NO HAY DATOS
352	SALTOS	FEMENINO	SOLTERO	MESTIZA	05/06/1987			-89	No existe dato	AMBATO	AMBATO	SI	NO HAY DATOS NO HAY DATOS
392	JURADO	FEMENINO	CASADO	MESTIZA	07/05/1986			-88	No existe dato	AMBATO	AMBATO	SI	NO HAY DATOS NO HAY DATOS
403	URRUTIA	FEMENINO	SOLTERO	MESTIZA	28/05/1984			-86	No existe dato	AMBATO	AMBATO	SI	NO HAY DATOS NO HAY DATOS
431	BUSTOS	FEMENINO	SOLTERO	MESTIZA	11/03/1988			-89	No existe dato	AMBATO	AMBATO	SI	NO HAY DATOS NO HAY DATOS

## 2.2. Bodega de datos

Posterior a retirar los datos atípicos se obtuvo una base de datos limpia con 425 estudiantes, se agregó la variable *Nivel*.

**Figura 4:** Bodega de datos

IDENTIFICADOR	GENERO	ESTADO CIVIL	ETNIA	EDAD	LUGAR NACI	CIUDAD RES	NIVEL	NOTA 1P	NOTA 2P	NOTA 1S	NOTA 2S	DESERTOR
1	MASCULINO	SOLTERO	MESTIZA	28	PELLEO	QUIRO	OCTAVO	8,2	6,9	7,5	7,1	NO
2	FEMENINO	SOLTERO	MESTIZA	25	AMBATO	QUIRO	DESCRNO	9,7	7,9	7,7	8,4	NO
3	FEMENINO	SOLTERO	MESTIZA	22	MÓCHA	MÓCHA	DESCRNO	8,1	7,1	7,5	7,7	NO
4	FEMENINO	CASADO	MESTIZA	19	PILLARO	PILLARO	SEXTO	8,6	8,5	8,7	8,7	SI
5	FEMENINO	SOLTERO	MESTIZA	22	PILLARO	PILLARO	OCTAVO	8,5	8,0	7,8	8,3	NO
6	FEMENINO	SOLTERO	MESTIZA	26	PILLARO	PILLARO	NOVENO	7,9	8,3	8,4	8,3	NO
7	MASCULINO	SOLTERO	MESTIZA	22	PILLARO	PILLARO	PRIMERO	8,4	8,2	8,8	8,0	SI
8	FEMENINO	SOLTERO	MESTIZA	19	PILLARO	PILLARO	OCTAVO	8,1	7,5	8,7	7,8	NO
9	MASCULINO	CASADO	MESTIZA	28	PILLARO	PILLARO	SEGUNDO	7,9	7,5	7,1	8,0	SI
10	FEMENINO	SOLTERO	MESTIZA	19	PILLARO	PILLARO	OCTAVO	8,0	8,0	8,1	7,8	NO
11	MASCULINO	SOLTERO	INDIGENA	28	PELLEO	PELLEO	DESCRNO	8,2	8,2	7,8	7,9	NO
12	MASCULINO	CASADO	INDIGENA	22	PELLEO	PELLEO	DESCRNO	7,6	7,5	6,5	6,6	NO
13	MASCULINO	SOLTERO	MESTIZA	21	PELLEO	PELLEO	OCTAVO	7,8	7,8	8,1	8,1	NO
14	FEMENINO	SOLTERO	MESTIZA	26	PELLEO	PELLEO	CUARTO	7,2	7,8	7,8	7,8	SI
15	MASCULINO	CASADO	MESTIZA	22	PELLEO	PELLEO	NOVENO	7,5	8,0	7,9	7,7	NO
16	FEMENINO	DIVORCIADO	MESTIZA	25	PELLEO	PELLEO	OCTAVO	8,5	8,4	7,4	7,8	NO
17	MASCULINO	SOLTERO	MESTIZA	19	PELLEO	PELLEO	SEGUNDO	7,4	7,1	6,5	6,6	SI
18	MASCULINO	SOLTERO	INDIGENA	19	PELLEO	PELLEO	PRIMERO	7,9	6,9	8,0	8,0	SI
19	MASCULINO	SOLTERO	INDIGENA	23	PELLEO	PELLEO	OCTAVO	8,2	8,2	7,8	8,1	NO
20	FEMENINO	SOLTERO	MESTIZA	28	AMBATO	PELLEO	CUARTO	8,0	8,2	8,2	8,0	SI
21	FEMENINO	SOLTERO	MESTIZA	19	AMBATO	PELLEO	DESCRNO	8,6	8,6	8,1	7,5	NO
22	FEMENINO	SOLTERO	MESTIZA	18	AMBATO	PELLEO	OCTAVO	8,9	7,9	7,5	7,5	NO
23	MASCULINO	SOLTERO	MESTIZA	23	AMBATO	PELLEO	DESCRNO	8,0	7,2	7,8	7,9	NO
24	MASCULINO	SOLTERO	INDIGENA	23	AMBATO	PELLEO	DESCRNO	8,6	6,7	7,2	7,4	NO

Fuente: Datos Históricos a partir del año 2016

## ETAPA III

### TRANSFORMACIÓN DE DATOS

#### 3.1. Selección de datos

Los datos puros provienen de la base de datos relacional de la Universidad Técnica de Ambato, se tomó *datos* en diferentes hojas de cálculo de 485 estudiantes de la Carrera de Docencia en Informática a partir del año 2006 los datos fueron transformados a datos abstractos según los tipos de los Atributos propuestos por el estadístico S. Stevens (1946). Se clasificaron en tres tipos:

- Nominal (los valores son básicamente una función de etiquetado,
- Ordinal (los valores obedecen a la relación de orden)
- Cuantitativo (los valores tienen todo el poder expresivo de los números reales

#### 3.2. Análisis de Variables

Tabla 74: Variables Predictivas

Atributos	Tipo Atributo S. Stevens (1946)
Genero	Cualitativa - Nominal
Estado Civil	Cualitativa - Nominal
Etnia	Cualitativa - Nominal
Edad	Cuantitativo
Lugar de Nacimiento	Cualitativa - Nominal
Ciudad de residencia	Cualitativa - Nominal
Nivel	Cualitativa – Ordinal
Nota1p	Cuantitativo
Nota2p	Cuantitativo
Nota1s	Cuantitativo
Nota2s	Cuantitativo

Elaborado por: Blanca Cuji

**Variable a predecir (Variable discriminante)**

Desertor	Cualitativa Nominal - variable a predecir
----------	---

**ETAPA IV**

**MINERÍA DE DATOS**

Puesto que el objetivo es generar un modelo predictivo, se toma como variable predictiva a “*Desertor*, cualitativa de tipo nominal.

Para la generación del modelo se tomó en cuenta el siguiente esquema:

**Esquema utilizado para la generación del Modelo predictivo**

**Tabla 75: Tabla de Aprendizaje (90 % de datos)**

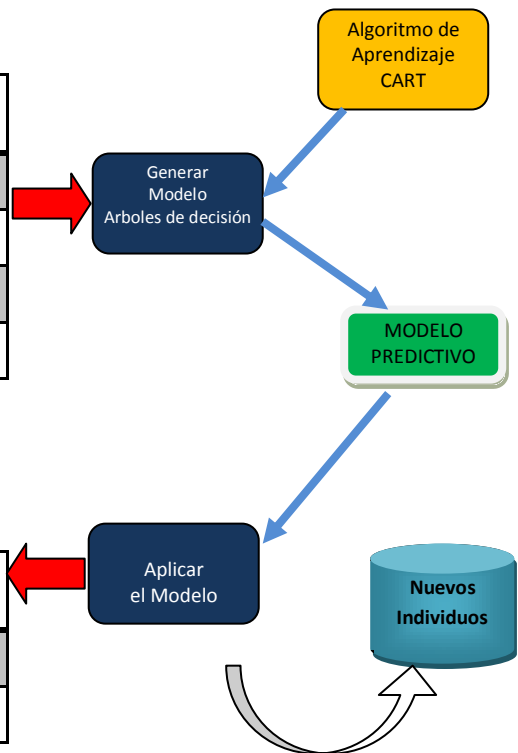
Id	Genero	Estado civil	Nota1p	Nota2p	Desertor
1	MASCULINO	SOLTERO	8,2	6,9	NO
2	FEMENINO	SOLTERO	9,7	7,9	NO
3	FEMENINO	SOLTERO	8,1	7,1	SI
4	FEMENINO	SOLTERO	5,1	7,1	SI

Elaborado por: Blanca Cuji

**Tabla 76: Tabla de Testing (10% de datos)**

Id	Genero	Estado civil	Nota1p	Nota2p	Desertor
1	MASCULINO	SOLTERO	8,2	6,9	NO
2	FEMENINO	SOLTERO	9,7	7,9	NO

Elaborado por: Blanca Cuji



#### 4.1. Selección de Algoritmo

Se utiliza el algoritmo de clasificación con *Arboles de decisión*. Para la generación del algoritmo se utiliza el método CART (Classification and regression tree – Arboles de clasificación), que utiliza el algoritmo “*top-down*” (Algoritmo Hunt) para construir el árbol de arriba hacia abajo.

Nota: Se usó árboles de decisión para generar el modelo porque se cuentan con variables cualitativas y cuantitativas en los datos, además el número de variables es apropiado para su aplicación.

#### Explicación del método CART (Algoritmo Hunt)

- Sea  $D_t$  (datos de entrenamiento) el conjunto de registros de entrenamiento en un nodo  $t$  dado.
- Sea  $y_t = \{y_1, y_2, \dots, y_c\}$  (datos asignados a una clase= Desertor ó No desertor) el conjunto de etiquetas de la clase.
- Si todos los registros  $D_t$  pertenecen a la misma clase  $y_t$  entonces  $t$  es un nodo hoja (nodo puro las variables pertenece a una misma clase).
- Si  $D_t$  contiene registros que pertenecen a más de una clase se escoge una variable (atributo) para dividir los datos en subconjuntos más pequeños.
- Se aplica este mismo proceso en forma recursiva a cada subconjunto de datos

#### Escoger el orden de las variables para construir el árbol de decisión

Nota: No se puede usar árboles de decisión en una gran cantidad de variables (más de 500 variables)

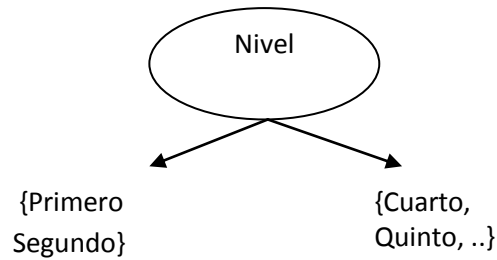
#### 1. ¿Cómo se dividirán las variables que se encuentran en la base de datos?

Con *divisiones binarias*, el método CART solo utiliza divisiones binarias así:

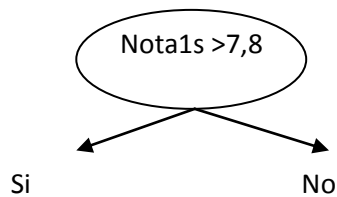
## Nominales



## Ordinales



## Cuantitativas: partir del punto medio



2. ¿Qué atributos se van a utilizar y en qué orden? ¿Qué criterios se debe utilizar para escoger la “mejor” división en el árbol?

## Pureza e impureza de los nodos

Para medir la pureza de un nodo se consideran los siguientes índices:

1. Error de Clasificación: (0- 0,5: 0= cuando es un nodo 100% puro y 0,5=cuando es 100% impuro)
2. Índice Gini: (0- 0,5: 0= cuando es un nodo 100% puro y 0,5=cuando es 100% impuro)
3. Entropía: (0- 1: 0= cuando es un nodo 100% puro y 1=cuando es 100%

## Fórmula:

$p(j/t)$  = Probabilidad de pertenecer a la clase “j” estando en el nodo t

donde:

p=probabilidad

j: clase (si –no)

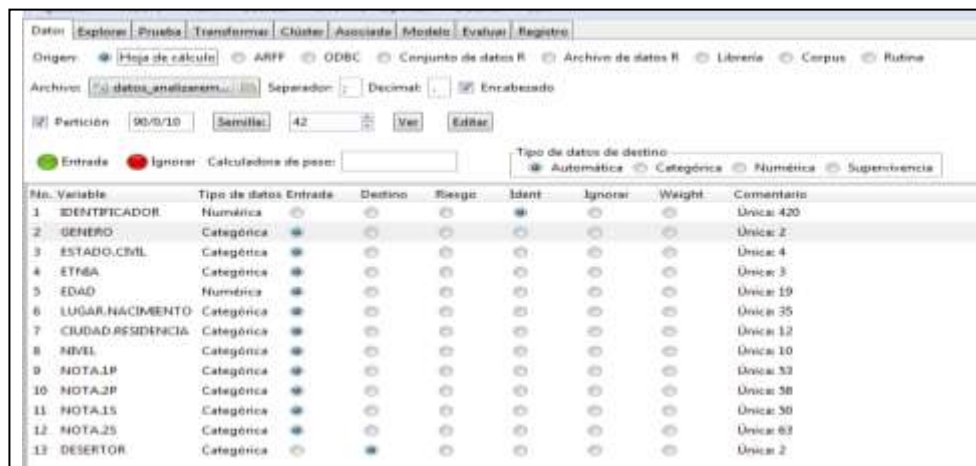
t= Nodo

Se escogió la mejor división basado en la menor variabilidad después de aplicar los índices propuestos.

### 4.3. Generación del Modelo

Para la generación del modelo predictivo, se usó Rattle de R; se construyó el árbol con los datos de los estudiantes de la Carrera de Docencia en Informática.

**Figura 5:** Variables para realizar el modelo



Nº. Variable	Tipo de datos	Entrada	Destino	Resq:	Ident	Ignorar	Weight	Comentario
1 IDENTIFICADOR	Numérica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 420
2 GENERO	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 2
3 ESTADO.CIVIL	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 4
4 ETNIA	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 3
5 EDAD	Numérica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 19
6 LUGAR.NACIMIENTO	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 25
7 CIUDAD.RESIDENCIA	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 12
8 NIVEL	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 10
9 NOTA.1P	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 33
10 NOTA.2P	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 38
11 NOTA.1S	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 30
12 NOTA.2S	Catagórica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 63
13 DESERTOR	Catagórica	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Única: 2

**Fuente:** Datos Históricos a partir del año 2016

Para la construcción del modelo se tomó el **90%** de los datos para la tabla de aprendizaje y el **10%** para la tabla de testing. Se generó el modelo con el algoritmo rpart de CART con 4 observaciones en cada nodo del árbol y cuatro niveles de profundidad.





## Reglas de decisión básicas obtenidas a partir del árbol de decisión

### Árbol como reglas:

Rule number: 3 [DESERTOR=SI cover=111 (29%) prob=1.00]  
NIVEL=CUARTO,PRIMERO,QUINTO,SEGUNDO,TERCERO

Rule number: 11 [DESERTOR=SI cover=5 (1%) prob=1.00]  
NIVEL= DÉCIMO,NOVENO,OCTAVO,SÉPTIMO,SEXTO  
NOTA.1S=7,2,7,4,7,9,8,7,9,1  
NOTA.2S=5,6,7,6,8,5,8,7

Rule number: 10 [DESERTOR=NO cover=33 (9%) prob=0.09]  
NIVEL=DÉCIMO,NOVENO,OCTAVO,SÉPTIMO,SEXTO  
NOTA.1S=7,2,7,4,7,9,8,7,9,1

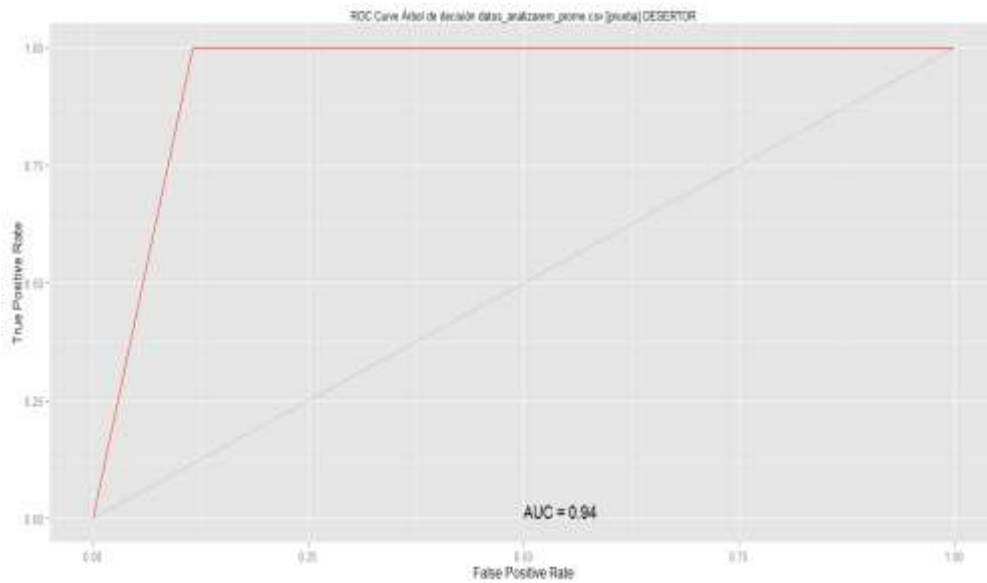
NOTA.2S=5,4,6,7,6,8,7,2,7,4,7,5,7,7,7,8,7,9,8,2,8,3,8,4,8,6,8,8,9,0,9,1,9,2,9,5,9,6

Rule number: 4 [DESERTOR=NO cover=229 (61%) prob=0.00]  
NIVEL=DÉCIMO,NOVENO,OCTAVO,SÉPTIMO,SEXTO  
NOTA.1S=4,0,4,5,6,0,6,3,6,5,6,6,6,8,6,9,7,0,7,1,7,3,7,5,7,6,7,7,7,8,8,0,8,1,8,2,8,  
3,8,4,8,5,8,6,8,8,8,9,9,0,9,2,9,3,9,4

### Evaluación del modelo

Se construyó la curva ROC para evaluar la efectividad del modelo

**Figura 7:** Evaluación Modelo



**Fuente:** Datos Históricos a partir del año 2016

Según la gráfica el modelo tiene una probabilidad de efectividad en la predicción del **94%**.

## ETAPA V

### INTERPRETACIÓN

#### 5.1. Visualización

Se diseñó una interfaz web con las variables utilizadas para la generación del modelo predictivo de deserción, se utilizó las reglas generadas por el modelo para predecir futuros desertores.

### Cumplimento Regla 3



**FACULTAD DE CIENCIAS HUMANAS  
Y DE LA EDUCACIÓN**



💡 Teclear los datos para un alumno

Cédula :	<input type="text" value="1803127594"/>		
Nombre :	<input type="text" value="Blanca"/>		
Apellido :	<input type="text" value="Cuji"/>		
Genero :	<input type="text" value="Femenino"/>	Estado Civil :	<input type="text" value="Casado(a)"/>
Etnia :	<input type="text" value="Mestizo"/>	Edad :	<input type="text" value="37"/>
Lugar Nacimiento :	<input type="text" value="Ambato"/>	Ciudad Residencia :	<input type="text" value="Ambato"/>
Nivel :	<input type="text" value="Primero"/>		
Nota IP :	<input type="text" value="0"/>	Nota 2P :	<input type="text"/>
Nota IS :	<input type="text"/>	Nota 2S :	<input type="text"/>

💡 RESULTADOS

Cumple Regla 3:

Probabilidad de Desertar :  Porcentaje de Desertar :  %



**FCHE**  
Dirección: Av. Los Chaguá y Río Cutuchi Ciudadela Universitaria

2016

### Reporte:

ID	CEDULA	NOMBRE	APELLIDO	GENERO	ESTADO_CIVIL	ETNIA	EDAD	LUGAR_NACIMIENTO	CIVIL	IP	IS	2P	2S	DESBOR	PROBABLE	REALIZADO
1	1803127594	BLANCA	CUJI	FEMENINO	CASADO(A)	MESTIZO	37	AMBATO	1	0	0	0	0	0	29	SI

1 de 1 Registros

## Cumplimiento Regla 4:



**FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACIÓN**

💡 Teclear los datos para un alumno

Cedula :	1803127594		
Nombre :	Susana		
Apellido :	Gonzalez		
Genero :	Femenino ▾	Estado Civil :	Casado(a) ▾
Etnia :	Indigena ▾	Edad :	20
Lugar Nacimiento :	Quito	Ciudad Residencia :	Quito
Nivel :	Septimo ▾		
Nota 1P :	7.00	Nota 2P :	7.00
Nota 1S :	8.00	Nota 2S :	7.00

💡 RESULTADOS

Cumple Regla 4:

Probabilidad de Desertar :  Porcentaje de Desertar :  %



**FCHE**  
Dirección: Av. Los Chasquis y Río Cutachi Ciudadela Universitaria

2016

## Reporte

DETENGA - Google Chrome

localhost/www/sisicf/historial.php

**HISTORIAL DE CALCULOS**

Buscar

Opciones:  Por Nombre

ID	CEDELA	NOMBRE	APELLIDO	GENERO	ESTADO_CIVIL	ETNIA	EDAD	LUGAR_NACIMIENTO	CIVIL	NOTAS1P	NOTAS1S	NOTAS2P	NOTAS2S	DESIERTO	PROBABILIDAD	FECHA REGISTRO
1	1803127594	REAYCA	CID	FEMENINO	CASADO(A)	INDIGENA	17	ALBAYDO	1	7	8	7	7	SI	20	2016-07-20 11:59:28
2	1803127594	SUSANA	GONZALEZ	FEMENINO	CASADO(A)	INDIGENA	20	QUITO	1	7	8	7	7	NO	61	2016-07-20 12:58:12

50

Página 1 de 1

Último 1 de 1 de 1 de 1



## FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACION



### HISTORIAL DE CALCULOS

ID	CEDELA	NOMBRE	APELLIDOS	GENERO	ESTADO_CIVIL	ETNIA	EDAD	LUGAR_NACIMIENTO	CUIDAD_RESIDENCIA	NIVEL	NOTA1P	NOTA2P	NOTA3P	NOTA4P	DESBERTOR	PROBABILIDAD	CREATED
1	288822784	ALANCA	CAB	FEMININO	CASADO	MESTIZO	37	AMBATO	AMBATO	3	7	8	7	3	61	24	2016-07-20 22:50:38
2	288822784	ISABELE	ESPINOZA	FEMININO	CASADO	MESTIZO	38	QUITO	QUITO	3	7	7	8	3	60	41	2016-07-20 22:56:32

Los datos de los estudiantes se almacenan en una base de datos realizada en Mysql en donde se almacenan las variables:

#	Nombre	Tipo	Caracteres	Atributos	Notas	Predefinidos	Extra	Acción
1	ID	INT(11)			No	Unsigned	AUTO_INCREMENT	Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
2	CEDELA	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
3	NOMBRE	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
4	APELLIDOS	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
5	GENERO	VARCHAR(10)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
6	ESTADO_CIVIL	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
7	ETNIA	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
8	EDAD	INT(11)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
9	LUGAR_NACIMIENTO	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
10	CUIDAD_RESIDENCIA	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
11	NIVEL	INT(11)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
12	NOTA1P	Real			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
13	NOTA2P	Real			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
14	NOTA3P	Real			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
15	NOTA4P	Real			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
16	DESBERTOR	VARCHAR(255)			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
17	PROBABILIDAD	Real			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
18	CREATED	datetime			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias
19	MODIFIED	datetime			SI			Cancelar Eliminar Privilegios Operaciones Seguridad Dependencias

### 6.8.1. Conclusiones de la propuesta

- La variable *Nivel* es aquella que mayor incide en la deserción. Se muestra que los estudiantes que se encuentran en Primero, Segundo, Tercero, cuarto y quinto nivel son aquellos que tienen mayor tendencia a desertar en niveles como Sexto y Séptimo la deserción es mínima y en Octavo, Noveno y Décimo la deserción es nula.
- Los estudiantes que se encuentren en el NIVEL=DECIMO,NOVENO,OCTAVO,SEPTIMO,SEXTO y que tienen notas comprendidas entre (4,0; 4,5; 6,0; 6,3; 6,5; 6,6; 6,8; 6,9; 7,0; 7,1; 7,3; 7,5; 7,6; 7,7; 7,8; 8,0; 8,1; 8,2; 8,3; 8,4; 8,5; 8,6; 8,8; 8,9; 9,0; 9,2; 9,3; 9,4) correspondiente a la variable Nota.1S no son desertores.
- De los 485 con 15475 registros analizados la variable edad y Nota.1s muestran datos atípicos (observación extrañamente grande o pequeña).
- El modelo propuesto es un punto de partida en la generación de modelos de predicción usando minería de datos, para analizar y evaluar los factores que influyen en la deserción universitaria.

### 6.8.2. Recomendaciones de la propuesta

- Específicamente, se debe hacer un seguimiento a los estudiantes de primer, segundo y tercer semestre que provienen de lugares como Zamora Chinchipe, Puyo, Latacunga, Cuenca, Guayaquil, Babahoyo en materias como Programación y física para disminuir el alto grado de deserción.
- Se recomienda alimentar la base de datos con información de los estudiantes a partir del año 1998 (año de creación de la Carrera), hasta el año 2005, pues solo se encuentra almacenada en archivos físicos lo que impide tener una mayor cantidad de datos; esto ayudará a dar mayor soporte a futuros proyectos encaminados a disminuir la deserción escolar.
- Aplicar otros métodos de clasificación para validar el método más adecuado que se adapte a la estructura de los datos suministrados por Dirección de Tecnologías de Información y Comunicaciones (DITCI) que posee la UTA.
- Profundizar en la categorización y determinación de las variables que poseen una mayor incidencia en el problema de la deserción y establecer un procedimiento de captura de los datos de dichas variables, en el momento que el alumno registra su matrícula en la Universidad.
- Variables recomendadas para mejorar el modelo en trabajos futuros

**Tabla 77: Variables recomendadas para trabajos futuros**

<b>Factores</b>	<b>Variable</b>
<b>Socioeconómico</b>	Nivel socioeconómico
	Nivel educativo de la Madre
	Nivel educativo del Padre
	Situación laboral del padre
	Trabajo del estudiante
	Financiamiento_estudios
	Horas de trabajo_estudiante
	Promedio_Ingresos_familiares
<b>Individuales</b>	Fecha de Nacimiento

	Lugar de Nacimiento
	Lugar de residencia
	Genero
	Vivienda
	Nucleo_familiar (vive con familiares u otros)
	Practica Deporte
	Hábitos de Consumo
	Nota de graduación del nivel medio
	Decisión_elección_carrera (propia )
	Estado civil
	Número de hijos
	Número de horas dedicadas al trabajo autónomo
<b>Académicos</b>	Nota que obtuvo al ingreso a la Universidad
	Rendimiento académico
	Fecha de Ingreso a la universidad
	Periodo Académico
	Asignatura
	Experiencia docente
	Especialidad docente por asignatura
	Duración de clases
	Métodos de evaluación del docente
<b>Institucionales</b>	Sede donde estudia el alumno
	Colegio publico
	Colegio privado
	Año de egreso de la secundaria
	Tipo de Bachillerato
<b>Culturales</b>	Nacionalidad
	Etnia

Fuente: (Márquez, 2015)



- Con base en las variables y factores propuestos en la *Tabla 78: Variables recomendadas para trabajos futuros* se puede derivar perfiles de posibles estudiantes desertores y construir modelos predictivos de mayor precisión sean estos matemáticos o con Data Mining.
- Con base en la investigación propuesta por (Argote, Jimenez, & Gòmez, 2014) en la cual se exponen cuatro modelos matemáticos predictivos basado en diferentes variables resultados de la investigación, se recomienda como propuesta futura la generación de modelos matemáticos similares a los que se exponen a continuación:

1. Modelo Matemático: Factor Socioeconómico

<b>Variables Nivel socioeconómico( NSE)</b>	<b>Denominación</b>
Nivel educativo de la Madre	NEM
Nivel educativo del Padre	NEP
Situación laboral del padre	SLP
Trabajo del estudiante	TE
Financiamiento_estudios	FES
Horas de trabajo_estudiante	HTE
Promedio_Ingresos_familiares	PIF

Modelo matemático basado en (Argote, Jimenez, & Gòmez, 2014)

$$NSE = \frac{\frac{NEM}{\%IE} + \frac{NEP}{\%INEP} + \frac{SLP}{\%ISLP} + \frac{TE}{\%ITE} + \frac{FES}{\%IFES} + \frac{HTE}{\%IHTE} + \frac{PIF}{\%PIF}}{7} * 100$$

2. Modelo Matemático: Factores Individuales

<b>Individuales (IND)</b>	<b>Denominación</b>
Fecha de Nacimiento	FN
Lugar de Nacimiento	LN
Lugar de Residencia	LR
Genero	G

Vivienda	V
Nucleo_familiar (vive con familiares u otros)	NF
Practica Deporte	PD
Hábitos de Consumo	HC
Nota de graduación del nivel medio	NG
Decisión_elección_carrera (propia )	DEC
Estado civil	EC
Número de hijos	NH
Número de horas dedicadas al trabajo autónomo	HTA

Modelo matemático basado en (Argote, Jimenez, & Gómez, 2014)

$$IND = \frac{\frac{FN}{\%IFN} + \frac{LN}{\%LN} + \frac{LR}{\%ILR} + \frac{G}{\%IG} + \frac{V}{\%IV} + \frac{NF}{\%INF} + \frac{DP}{\%IDP} + \frac{HC}{\%IHC} + \frac{NG}{\%ING} + \frac{DEC}{\%IDEC} + \frac{EC}{\%IEC} + \frac{NH}{\%INH} + \frac{HTA}{\%IHTA}}{13} * 100$$

### 3. Modelo Matemático: Factores Institucionales

Institucionales (INS)	Denominación
Sede donde estudia el alumno	SU
Colegio publico	CPU
Colegio privado	CPI
Año de egreso de la secundaria	AES
Tipo de Bachillerato	TB

Modelo matemático basado en (Argote, Jimenez, & Gómez, 2014)

$$INS = \frac{\frac{SU}{\%ISU} + \frac{CPU}{\%ICPU} + \frac{CPI}{\%ICPI} + \frac{AES}{\%IAES} + \frac{TB}{\%ITB}}{5} * 100$$

#### 4. Modelo Matemático: Factores Culturales

<b>Culturales (C)</b>	<b>Denominación</b>
<b>Nacionalidad</b>	N
<b>Etnia</b>	ET

Modelo matemático basado en (Argote, Jimenez, & Gómez, 2014)

$$C = \frac{\frac{N}{\%IN} + \frac{ET}{\%IET}}{2} * 100$$

## BIBLIOGRAFÍA

Aldehuela, M. (2015). Análisis comparativo entre Métodos estadísticos y de Minería de Datos. *ICAI-Universidad Pontificia Comillas* .

Apaza, E. (2012). Factores determinantes que inciden en la deserción de los estudiantes universitarios. *Apuntes Universitarios* , 77-86.

Argote, I., Jimenez, R., & Gómez, J. (2014). DETECCIÓN DE PATRONES DE DESERCIÓN EN LOS PROGRAMAS DE PREGRADO DE LA UNIVERSIDAD MARIANA DE SAN JUAN DE PASTO, APLICANDO EL PROCESO DE DESCUBRIMIENTO DE CONOCIMIENTO SOBRE BASE DE DATOS (KDD) Y SU IMPLEMENTACIÓN EN MODELOS MATEMÁTICOS DE PREDICCIÓN. *IV Conferencia Latinoamericana sobre el abandono en la Educación Superior*, (págs. 1-7). Colombia.

BRITOS, P., & HOSSIAN, A. (2005). *Minería de Datos*. Argentina: Nueva Librería.

Caridad, J. M. (2011). *Minería de Datos: Análisis de Bases de Datos en la empresas*. España.

Consejo Nacional de Planificación. (19 de Octubre de 2013). Plan Nacional del Buen Vivir. *Plan Nacional del Buen Vivir* . Quito, Pichincha, Ecuador.

Diario Hoy. (8 de Noviembre de 2008). La deserción universitaria amenaza la gratuidad de la educación pública. *La deserción universitaria amenaza la gratuidad de la educación pública* .

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

Essax, A., & Ayad, H. (2012). Improving student success using predictive models and data visualisations. *Learning Technology* , 58-70.

Estellés, J., Alcover, R., Dapena-Janeiro, A., Valderruten, A., Satorre, F., & Llopis, F. (2009). *Rendimiento Académico de los Estudios de Informática en Algunos Centros Españoles*. Barcelona: XV Jornadas de Enseñanza Universitaria de la Informática.

Gerben, W. D., Mykola, P., & Jan, M. V. (2009). Predicting Students Drop Out: A Case Study. *Educational Data Mining* , 41-50.

González González, M. (2006). Absentismo y Abandono Escolar: Una situación Singular de la Exclusión Educativa. *Revista electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación* , Vol. 4. 1 - 15.

Hope, B. (2016). La minería de datos compite con el corredor de bolsa. *ProQuest* , 3-4.

Jadrić, M., Garača, Ž., & Željko, M. (2010). STUDENT DROPOUT ANALYSIS WITH APPLICATION OF DATA MINING METHODS . *Preliminary communication* , 31-46.

KIMBALL, R. (2002). *The data warehouse toolkit: the complete guide to dimensional*. USA, ISBN: 780471200246: Wiley Computer Publishing.

Márquez, C. (2015). *Predicción del fracaso y el abandono escolar mediante técnicas de Minería de Datos*. Córdoba: Publicaciones de la Universidad de Córdoba.

Marquez, C., & Romero, C. (2011). Predicting School Failure Using Data Mining. *IEEE Explorer* , 1-5.

Martínez, J., & Pérez, J. (2008). Efecto de la Trayectoria Académica en el Desempeño de Estudiantes de Ingeniería en Evaluaciones Nacionales. *Formación Universitaria* , 3-12.

Medina, S. (2011). *Minería de datos*. Colombia.

MORALES, E. (11 de Julio de 2009). <http://ccc.inaoep.mx/~emorales/Cursos/KDD/principal.html>. Obtenido de <http://ccc.inaoep.mx/~emorales/Cursos/KDD/principal.html>.

Rubio, M., Tocaín, A., & Mantila, M. (05 de julio de 2012). *www.pucesi.edu.ec*. Recuperado el 11 de 2015, de Pontificia Universidad Católica sede Ibarra: <http://www.pucesi.edu.ec/proyectoalfa/files/final2012.pdf>

Salazar, A., Gosálbez, J., Miralles, R., & Vergara, L. (2004). A case study of knowledge discovery on academic achievement, student desertion and student retention. *IEEE Xplore* , 150-154.

Santamaria, W. (2015). *researchgate*. Recuperado el Enero de 2016, de [researchgate.net](http://researchgate.net):

[https://www.researchgate.net/publication/240724702\\_Tecnicas\\_de\\_Mineria\\_de\\_Datos\\_Aplicadas\\_en\\_la\\_Deteccion\\_de\\_FraudeEstado\\_del\\_Arte](https://www.researchgate.net/publication/240724702_Tecnicas_de_Mineria_de_Datos_Aplicadas_en_la_Deteccion_de_FraudeEstado_del_Arte)

Timarán, R., & Calderón, A. (2013). La minería de datos como método innovador para la detección de patrones de deserción estudiantil en programas de pregrado en Instituciones de Educación Superior. *WEEF* , 2-9.

Tutiempo Network. (12 de junio de 2015). <http://www.tutiempo.net/Modelos/>.  
Obtenido de <http://www.tutiempo.net/Modelos/>:  
<http://www.tutiempo.net/Modelos/>

Valero, S., Salvador, A., & García, M. (2008). Minería de datos: predicción de la deserción escolar. *Universidad Tecnológica de Izúcar de Matamoros mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos* , 8.

Villena, J., & Crespo, R. M. (2012). Minería de datos. *Data Mining. DAEDALUS White Paper* , 2-39.

Zhang, G. (2000). Neural Networks for Classification. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART C:APPLICATIONS AND REVIEWS* , Vol. 30, págs. 451 - 462.

Zhao, Q., & Bhowmick, S. (2006). Association Rule Mining: A Survey. *Nanyang Technological University* .

## ANEXOS

### Anexo 1: Formato de encuesta

#### TÉCNICAS DE PREDICCIÓN Y DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL

**Objetivo:** Detectar factores que inciden en la Deserción Estudiantil

**Destinatarios:** Estudiantes de la Carrera de Docencia en Informática

**Instrucciones:**

Marque con una X en el casillero que corresponda.

Los datos obtenidos serán confidenciales y se usarán con efectos investigativos.

#### I. DATOS GENERALES:

<b>Edad</b>	16-20 años	21-25 años	26-35 años	Más de 35	
<b>Lugar de Nacimiento</b>					
<b>Genero</b>	Hombre		Mujer		
<b>Estado Civil</b>	Casado	Soltero	Divorciado	Unión Libre	Viudo
<b>Capacidad Especial</b> Si No	Visual	Física	Psicológica	Intelectual	Sensorial
<b>Identificación Étnica</b>	Mestizo/a	Blanco/a	Indígena	Afrodescendiente	

#### II. ENCUESTA

1. ¿Cuál de las siguientes razones propician la Deserción Estudiantil?
  - Dificultades económicas
  - Sanción Disciplinaria
  - Aplazamiento de Semestre
  - Bajo rendimiento académico
  - Cambio de carrera
  - Cambio de ciudad
  - Dificultades familiares
  - Enfermedad
  - Ubicación laboral
  - Otros
2. La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos:
  - Humanos
  - Materiales
  - Tecnológicos
  - Otros  Cual? .....
3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?
 

Si  No
4. Al culminar el semestre Octubre 2015 – Marzo 2016, usted cree que:
  - Aprobará todas sus materias con buenas calificaciones.
  - Aprobará todas sus materias con calificaciones regulares.
  - Probablemente repruebe una materia.
  - No creo aprobar este semestre
5. ¿Cuánto tiempo dedica a sus estudios como parte de su trabajo autónomo (fuera de clases) ?
  - Menos de 1 h.
  - De 1 a 2 h.
  - Más de 2 h.
6. La decisión de estudiar la carrera universitaria en su momento...
  - La tome yo.
  - Mis padres y/o familiares influyeron en mi decisión.

7. ¿Ha tenido alguna enfermedad grave?  
 Si  Cual ..... No
8. En cuanto al nivel económico de ingreso en su familia es...  
 • Alto.   
 • Medio.   
 • Bajo.   
 • Muy bajo.
9. ¿Dispone de los recursos económicos para costear todos los gastos que se necesita para estudiar?  
 Si  No
10. ¿Cuenta con algún tipo de ayuda o beca para realizar sus estudios?  
 Si  No
11. ¿Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia?  
 • 10%   
 • 20%   
 • 30%   
 • 40%   
 • 50%   
 • 60%   
 • 70%   
 • 80%   
 • No conoce
12. Los estudiantes abandonan sus estudios cuando cursan el..... de sus créditos  
 • 10%   
 • 20%   
 • 30%   
 • 40%   
 • 50%   
 • 60%   
 • 70%   
 • 80%   
 • No conoce
13. Del nivel de estudios de sus padres. Marca con una X:  

<b>MADRE</b>		<b>PADRE</b>	
Primaria	<input type="checkbox"/>	Primaria	<input type="checkbox"/>
Secundaria	<input type="checkbox"/>	Secundaria	<input type="checkbox"/>
Superior	<input type="checkbox"/>	Superior	<input type="checkbox"/>
Postgrado	<input type="checkbox"/>	Postgrado	<input type="checkbox"/>
No sé	<input type="checkbox"/>	No sé	<input type="checkbox"/>
14. ¿En su casa cuenta con un espacio adecuado (con mesa o escritorio) para estudiar?  
 Si  No
15. ¿Sus padres, fomentan y estimulan sus estudios?  
 Si  No
16. ¿Considera que los conocimientos que está adquiriendo en la institución son de utilidad para la vida?  
 • Sí.   
 • Sólo algunos.   
 • No.
17. ¿Es alguna asignatura especialmente difícil para Usted?  
 Si  Cual..... No
18. En el examen de admisión de ingreso a la universidad que nota obtuvo:  
 • Excelente.   
 • Muy bueno.   
 • Bueno.   
 • Regular.   
 • Malo.   
 • No presenté.



19. ¿Ha reprobado alguna asignatura de manera que se haya retrasado en algún periodo académico?  
Si  No

20. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?  
Si  No

Si su respuesta fue positiva escoja cuál de estos factores influyeron en su decisión

- Económicos y rendimiento académico
- Sociales ( migración, religión, nivel económico)
- Demográficos (edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)

## TÉCNICAS DE PREDICCIÓN Y DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL

**Objetivo:** Detectar factores que inciden en la Deserción Estudiantil

**Destinatarios:** Estudiantes desertores de la Carrera de Docencia en Informática

**Instrucciones:**

Marque con una X en el casillero que corresponda.

Los datos obtenidos serán confidenciales y se usarán con efectos investigativos.

**I. DATOS GENERALES:**

<b>Edad</b>	16-20 años	21-25 años	26-35 años	Más de 35	
<b>Lugar de Nacimiento</b>					
<b>Genero</b>	Hombre		Mujer		
<b>Estado Civil</b>	Casado	Soltero	Divorciado	Unión Libre	Viudo
<b>Capacidad Especial</b>	Visual	Física	Psicológica	Intelectual	Sensorial
<b>Si No</b>					
<b>Identificación Étnica</b>	Mestizo/a	Blanco/a	Indígena	Afrodescendiente	

**i. ENCUESTA**

1. ¿Cuál de las siguientes razones propician la Deserción Estudiantil?
  - Dificultades económicas
  - Sanción Disciplinaria
  - Aplazamiento de Semestre
  - Bajo rendimiento académico
  - Cambio de carrera
  - Cambio de ciudad
  - Dificultades familiares
  - Enfermedad
  - Ubicación laboral
  - Otros
2. La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos:
  - Humanos
  - Materiales
  - Tecnológicos

Otros Cual? .....
3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?
 

Si  No
4. En qué año abandono sus estudios en la Carrera de Docencia
  - 2010
  - 2011
  - 2012
  - 2013
  - 2014
  - Otro
5. ¿Cuánto tiempo dedicaba a sus estudios como parte de tu trabajo autónomo (fuera de clases) ?
  - Menos de 1 h.
  - De 1 a 2 h.
  - Más de 2 h.
6. La decisión de estudiar la carrera universitaria en su momento...
  - La tome yo.
  - Mis padres y/o familiares influyeron en mi decisión.
7. ¿Tiene o ha tenido alguna enfermedad grave?
 

Si  Cual ..... No
8. En cuanto al nivel económico de ingreso en su familia es...
  - Alto.
  - Medio.

- Bajo.
  - Muy bajo.
9. ¿Disponía de los recursos económicos para costear todos los gastos que se necesita para estudiar?  
Si  No
10. ¿Conto con algún tipo de ayuda o beca para realizar sus estudios?  
Si  No
11. Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia en Informática?
- 10%
  - 20%
  - 30%
  - 40%
  - 50%
  - 60%
  - 70%
  - 80%
  - No conoce
12. Los estudiantes abandonan sus estudios cuando cursan el..... de sus créditos
- 10%
  - 20%
  - 30%
  - 40%
  - 50%
  - 60%
  - 70%
  - 80%
  - No conoce
13. Del nivel de estudios de sus padres. Marca con una X:
- | <b>MADRE</b> |                          | <b>PADRE</b> |                          |
|--------------|--------------------------|--------------|--------------------------|
| Primaria     | <input type="checkbox"/> | Primaria     | <input type="checkbox"/> |
| Secundaria   | <input type="checkbox"/> | Secundaria   | <input type="checkbox"/> |
| Superior     | <input type="checkbox"/> | Superior     | <input type="checkbox"/> |
| Postgrado    | <input type="checkbox"/> | Postgrado    | <input type="checkbox"/> |
| No sé        | <input type="checkbox"/> | No sé        | <input type="checkbox"/> |
14. ¿En su casa contaba con un espacio adecuado (con mesa o escritorio) para estudiar?  
Si  No
15. ¿Sus padres, fomentaban y estimulan sus estudios?  
Si  No
16. ¿Considera que los conocimientos que adquirió durante su permanencia en la Carrera fueron de utilidad para la vida?
- Sí.
  - Sólo algunos.
  - No.
17. ¿Fue alguna asignatura especialmente difícil para Usted?  
Si  Cual..... No
18. En el examen de admisión de ingreso a la universidad que nota obtuvo:
- Excelente.
  - Muy bueno.
  - Bueno.
  - Regular.
  - Malo.
  - No presenté.
19. ¿Durante su permanencia en la Carrera reprobó alguna asignatura de manera que haya retrasado algún periodo académico?  
Si  No
20. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?  
Si  No

Si su respuesta fue positiva escoja cuál de estos factores influyeron en su decisión

- Económicos o rendimiento académico

- Sociales ( migración, religión, nivel económico)
- Demográficos (edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)



## TÉCNICAS DE PREDICCIÓN Y DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL

**Objetivo:** Detectar factores que inciden en la Deserción Estudiantil

**Destinatarios:** Egresados de la Carrera de Docencia en Informática

**Instrucciones:**

Marque con una X en el casillero que corresponda.

Los datos obtenidos serán confidenciales y se usaran con efectos investigativos.

### I- DATOS GENERALES:

<b>Edad</b>	16-20 años	21-25 años	26-35 años	Más de 35	
<b>Lugar de Nacimiento</b>					
<b>Genero</b>	Hombre		Mujer		
<b>Estado Civil</b>	Casado	Soltero	Divorciado	Unión Libre	Viudo
<b>Capacidad Especial</b>	Visual	Física	Psicológica	Intelectual	Sensorial
<b>Si No</b>					
<b>Identificación Étnica</b>	Mestizo/a	Blanco/a	Indígena	Afrodescendiente	

### II. ENCUESTA

1. Cuál de las siguientes razones propician la Deserción Estudiantil?
  - Dificultades económicas
  - Sanción Disciplinaria
  - Aplazamiento de Semestre
  - Bajo rendimiento académico
  - Cambio de carrera
  - Cambio de ciudad
  - Dificultades familiares
  - Enfermedad
  - Ubicación laboral
  - Otros
2. La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos:
  - Humanos
  - Materiales
  - Tecnológicos

Otros Cual? .....
3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?
 

Si  No
4. ¿Cuánto tiempo dedica a sus estudios como parte de su trabajo autónomo (fuera de clases) ?
  - Menos de 1 h.
  - De 1 a 2 h.
  - Más de 2 h.
5. La decisión de estudiar la carrera universitaria en su momento...
  - La tome yo.
  - Mis padres y/o familiares influyeron en mi decisión.
6. ¿Ha tenido alguna enfermedad grave?
 

Si  Cual ..... No
7. En cuanto al nivel económico de ingreso en su familia es...
  - Alto.
  - Medio.
  - Bajo.
  - Muy bajo.
8. ¿Disponía de los recursos económicos para costear todos los gastos que se necesita para estudiar?
 

Si  No
9. ¿Conto con algún tipo de ayuda o beca para realizar sus estudios?
 

Si  No

10. Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia?

- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- No conoce

11. Los estudiantes abandonan sus estudios cuando cursan el..... de sus créditos

- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- No conoce

12. Del nivel de estudios de tus padres. Marca con una X:

<b>MADRE</b>		<b>PADRE</b>	
Primaria	<input type="checkbox"/>	Primaria	<input type="checkbox"/>
Secundaria	<input type="checkbox"/>	Secundaria	<input type="checkbox"/>
Superior	<input type="checkbox"/>	Superior	<input type="checkbox"/>
Postgrado	<input type="checkbox"/>	Postgrado	<input type="checkbox"/>
No sé	<input type="checkbox"/>	No sé	<input type="checkbox"/>

13. ¿En su casa contaba con un espacio adecuado (con mesa o escritorio) para estudiar?

Si  No

14. ¿Sus padres, fomentaban y estimulaban sus estudios?

Si  No

15. ¿Considera que los conocimientos que adquirió en la institución fueron de utilidad para la vida?

- Si.
- Sólo algunos.
- No.

16. ¿Fue alguna asignatura especialmente difícil para Usted?

Si  Cual..... No

17. En el examen de admisión de ingreso a la universidad que nota obtuvo:

- Excelente.
- Muy bueno.
- Bueno.
- Regular.
- Malo.
- No presenté.

18. ¿Durante su permanencia en la Carrera reprobó alguna asignatura de manera que se haya retrasado en algún periodo académico?

Si  No

19. ¿Abandonado en alguna ocasión sus estudios en una institución educativa?

Si  No

Si su respuesta fue positiva escoja cuál de estos factores influyeron en su decisión

- Económicos
- Sociales ( migración, religión, nivel económico)
- Demográficos (edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)

## TÉCNICAS DE PREDICCIÓN Y LA DETECCIÓN DE PATRONES DE DESERCIÓN ESTUDIANTIL

**Objetivo:** Detectar factores que inciden en la Deserción Estudiantil

**Destinatarios:** Docentes de la Carrera de Docencia en Informática

**Instrucciones:**

Marque con una X en el casillero que corresponda.

Los datos obtenidos serán confidenciales y se usaran con efectos investigativos.

**I. DATOS GENERALES:**

<b>Edad</b>	16-20 años	21-25 años	26-35 años	Más de 35
-------------	------------	------------	------------	-----------

**II. ENCUESTA**

1. ¿Cuál de las siguientes razones propician la Deserción Estudiantil?
  - Dificultades económicas
  - Sanción Disciplinaria
  - Aplazamiento de Semestre
  - Bajo rendimiento académico
  - Cambio de carrera
  - Cambio de ciudad
  - Dificultades familiares
  - Enfermedad
  - Ubicación laboral
  - Otros
2. La automatización de procesos en los estudios sobre Deserción Estudiantil optimizará el uso de recursos:
  - Humanos
  - Materiales
  - Tecnológicos

Otros Cual? .....
3. ¿La técnica de predicción de minería de datos ayudará a detectar patrones de deserción escolar?
 

Si  No
4. Con respecto a sus estudiantes, al culminar el semestre Octubre 2015 – Marzo 2016 :
  - Aprobaran todas sus materias con buenas calificaciones.
  - Aprobarán todas sus materias con calificaciones regulares.
  - Probablemente reprobren alguna materia.
  - No creo que aprueben este semestre
5. ¿Cuánto tiempo considera que los estudiantes, dedican a sus estudios como parte de su trabajo autónomo (fuera de clases) ?
  - Menos de 1 h.
  - De 1 a 2 h.
  - Más de 2 h.
6. En cuanto al nivel económico de sus estudiantes es:
  - Alto.
  - Medio.
  - Bajo.
  - Muy bajo.
7. ¿Considera que los estudiantes disponen de los recursos económicos para costear todos los gastos que se necesita para estudiar?
 

Si  No
8. ¿Sus estudiantes cuentan con algún tipo de ayuda o beca para realizar sus estudios?
 

Si  No
9. Según su apreciación en que porcentaje considera que se da la Deserción Estudiantil en la Carrera de Docencia?
  - 10%
  - 20%
  - 30%
  - 40%
  - 50%
  - 60%

- 70%
  - 80%
  - No conoce
10. Los estudiantes abandonan sus estudios cuando cursan el..... de sus créditos
- 10%
  - 20%
  - 30%
  - 40%
  - 50%
  - 60%
  - 70%
  - 80%
  - No conoce
11. ¿Considera que los conocimientos que han adquirido los estudiantes en la institución son de utilidad para la vida?
- Sí.
  - Sólo algunos.
  - No.
12. ¿En su práctica docente ha tenido estudiantes que han desertado?
- Sí  No
- Si su respuesta fue positiva escoja cuál de estos factores considera que influyeron en su decisión
- Económicos y rendimiento académico
  - Sociales ( migración, religión, nivel económico)
  - Demográficos (edad, sexo, estado civil, trabajo, religión, lugar de residencia, etnia)
  - Rendimiento Académico