



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA
La Universidad Católica de Loja

ÁREA TÉCNICA

TÍTULO DE INGENIERO EN INFORMÁTICA

**Análisis de los usos de la tecnología en los
estudiantes de la Escuela Superior Politécnica de Chimborazo.**

TRABAJO DE TITULACIÓN.

AUTOR: Garcés Hernández, David Rafael

DIRECTORA: Jara Roa, Dunia Inés, Msg

CENTRO UNIVERSITARIO RIOBAMBA

2016

APROBACIÓN DE LA DIRECTORA DEL TRABAJO DE TITULACIÓN

Magister.

Dunia Inés Jara Roa.

DOCENTE DE LA TITULACIÓN

De mi consideración:

El presente trabajo de titulación: Análisis de los usos de la tecnología en los estudiantes de la Escuela Superior Politécnica de Chimborazo, realizado por Garcés Hernández David Rafael, ha sido orientado y revisado durante su ejecución, por lo que se aprueba la presentación del mismo.

Loja, noviembre de 2016

f)

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS

Yo, David Rafael Garcés Hernández, declaro ser autor del presente trabajo de titulación: Análisis de los usos de la tecnología en los estudiantes de la Escuela Superior Politécnica de Chimborazo, de la Titulación de Ingeniería en Informática, siendo la Magister Dunia Inés Jara Roa, directora del presente trabajo; y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales. Además, certifico que las ideas, conceptos, procedimientos y resultados vertidos en el presente trabajo investigativo, son de mi exclusiva responsabilidad.

Adicionalmente declaro conocer y aceptar la disposición del Art. 88 del Estatuto Orgánico de la Universidad Técnica Particular de Loja, que en su parte pertinente textualmente dice: “Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado que se realicen a través, o con el apoyo financiero, académico o institucional (operativo) de la Universidad”.

f.....

David Rafael Garcés Hernández

CI: 060250846-7

DEDICATORIA

A mis padres Fausto y Cecilia, ejemplo de trabajo y perseverancia; a mis hijas Dafne y Dulce fuente de inspiración para lograr todos mis objetivos.

David R. Garcés H.

AGRADECIMIENTO

A Dios, por darme fortaleza para cumplir mis sueños; a mi familia, quienes son el apoyo incondicional en mis aciertos y desaciertos; y a todas las personas que me apoyaron en este largo camino.

Especialmente a mi madre, por su empuje y motivación constante para cumplir este reto.

David R. Garcés H.

ÍNDICE DE CONTENIDOS

CARATULA	i
APROBACIÓN DE LA DIRECTORA DEL TRABAJO DE TITULACIÓN	ii
DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS	iii
DEDICATORIA	iv
AGRADECIMIENTO	v
ÍNDICE DE CONTENIDOS	vi
ÍNDICE DE TABLAS	x
ÍNDICE DE FIGURAS	xi
TERMINOLOGÍA	xii
ÍNDICE DE ANEXOS	xiii
RESUMEN	1
ABSTRACT	2
CAPITULO I	3
PLANTEAMIENTO DE LA INVESTIGACIÓN	3
1. Introducción.....	4
1.1. Problemática.....	5
1.2. Objetivos.....	6
1.2.1. General	6
1.2.2. Específicos.....	6
1.3. Preguntas de investigación.....	6
1.4. Hipótesis	6
CAPITULO II	7
MARCO TEÓRICO Y CONCEPTUAL	7
2. Marco teórico y conceptual	8
2.1. Brecha digital	8
2.2. Tipos de brecha digital.....	8

Brecha digital de acceso	8
2.3. Factores de la brecha digital.....	9
2.4. Internet.....	11
2.5. Uso del Internet en el Ecuador	12
2.6. Tecnología y educación.....	13
2.7. Áreas de desarrollo prioritario para la integración de las nuevas tecnologías en la educación.....	13
2.8. Nuevas prácticas educativas	13
2.9. Evaluación de aprendizajes.....	14
2.10. Educación virtual.....	14
2.10.1. E-learning.	15
2.10.2. B-learning.	15
2.10.3. M-learning.....	16
2.11. Las redes sociales en la educación	17
2.12. Mundos virtuales en la educación.....	18
2.13. Aproximación al concepto de minería de datos	18
2.14. Proceso de extracción del conocimiento a partir de datos (KDD).....	19
2.14.1. Selección de los datos.....	20
2.14.2. Limpieza de los datos.....	20
2.14.3. Transformación de los datos.	21
2.14.4. Minería de datos.....	21
2.14.5. Evaluación de patrones.....	21
2.14.6. Presentación del conocimiento.....	21
2.15. Fuentes de datos	23
2.16. Tipos de modelos de datos	24
2.16.1. Modelo descriptivo y sus operaciones.....	24
2.16.2. Modelo predictivo y sus operaciones.	24
2.17. Técnicas de minería de datos	25
2.17.1. Técnicas predictivas.....	25

2.17.2.	Técnicas descriptivas.	26
2.17.3.	Algoritmo simple k-means.	27
2.17.4.	Técnicas auxiliares.	28
2.18.	Sistemas de minería de datos.....	28
2.19.	Software estadístico y data mining	29
CAPITULO III		32
METODOLOGÍA		32
3.	Generalidades de la investigación	33
3.1.	Proceso de extracción del conocimiento a partir de datos (KDD)	33
3.2.	Etapa de selección de datos.....	34
3.2.1.	Fuentes de información.....	34
3.2.2.	Determinación de población y muestra.....	34
3.2.3.	Obtención de la información	35
3.3.	Etapa de exploración	36
3.4.	Etapa de limpieza y transformación.....	39
3.5.	Etapa de minería de datos.....	39
3.6.	Etapa de evaluación y presentación de resultados	39
CAPITULO IV.....		40
RESULTADOS.....		40
4.	RESULTADOS DE LA INVESTIGACIÓN	41
4.1.	Distribución de la muestra	41
4.2.	Datos Generales.....	41
4.2.1.	Genero	41
4.2.2.	Edad.....	42
4.2.3.	Ingresos	43
4.3.	Conocimiento y uso del Internet	44
4.3.1.	Preferencias de uso de Internet	44
4.3.2.	Internet para el aprendizaje	47
4.3.3.	Internet para el entretenimiento	51

4.4.	Rendimiento académico	54
4.5.	Perfiles de los estudiantes.....	55
4.5.1.	Perfil de uso del Internet para el aprendizaje.....	55
4.5.2.	Perfil de uso del Internet para entretenimiento.	56
4.6.	Pruebas de hipótesis	58
4.6.1.	Hipótesis 1.	58
4.6.2.	Hipótesis 2.	60
4.6.3.	Hipótesis 3.	63
4.6.4.	Hipótesis 4.	65
CAPITULO V.....		67
DISCUSIÓN DE RESULTADOS.....		67
5.	Discusión de resultados	68
5.1.	Perfil de los estudiantes.....	68
5.2.	Usos del Internet.....	70
5.3.	Discusión de las hipótesis	72
CONCLUSIONES		73
RECOMENDACIONES.....		74
TRABAJOS FUTUROS.....		75
REFERENCIAS		76
ANEXOS.....		80

ÍNDICE DE TABLAS

Tabla 1: Distribución de la muestra	36
Tabla 2: Características sociodemográficas	36
Tabla 3: Uso de Internet	36
Tabla 4: Uso Internet para aprendizaje	36
Tabla 5: Uso Internet entretenimiento	37
Tabla 6: Interacción en Internet	37
Tabla 7: Dispositivos para conectarse a Internet.....	37
Tabla 8: Internet como ayuda en lo académico.....	38
Tabla 9: Herramientas e-learning	38
Tabla 10: Rendimiento académico	38
Tabla 11: Uso de dispositivos.....	47
Tabla 12 : Exactitud de clasificación.....	56
Tabla 13: Número de casos en cada conglomerado	56
Tabla 14: Exactitud de clasificación.....	57
Tabla 15: Número de casos en cada conglomerado	57
Tabla 16: Tabla de clasificación	58
Tabla 17: Variables que no están en la ecuación.....	59
Tabla 18: Pruebas ómnibus sobre los coeficientes del modelo.....	59
Tabla 19: Prueba de Hosmer y Lemeshow	59
Tabla 20: Variables en la ecuación.....	60
Tabla 21: Tabla de clasificación	61
Tabla 22: Variables que no están en la ecuación.....	61
Tabla 23: Pruebas Ómnibus sobre los coeficientes del modelo.....	61
Tabla 24: Prueba de Hosmer y Lemeshow	62
Tabla 25: Variables en la ecuación.....	62
Tabla 26: Tabla de clasificación	63
Tabla 27: Variables que no están en la ecuación.....	64
Tabla 28: Pruebas Ómnibus sobre los coeficientes del modelo.....	64
Tabla 29: Variables en la ecuación.....	64
Tabla 30: Tabla de clasificación	65
Tabla 31: Variables que no están en la ecuación.....	66
Tabla 32: Pruebas omnibus sobre los coeficientes del modelo.....	66
Tabla 33: Variables en la ecuación.....	66

ÍNDICE DE FIGURAS

Figura 1: Uso del Internet en el Ecuador	12
Figura 2: Relación entre dato, información y conocimiento.....	19
Figura 3: Minería de Datos	19
Figura 4: Etapas del KDD	20
Figura 5: Resumen fases KDD	22
Figura 6: Sistemas de minería de datos	29
Figura 7: Logo SPSS	29
Figura 8: Logotipo Weka.....	30
Figura 9: Logotipo JHepWork	30
Figura 10: Logo RapidMiner	30
Figura 11: Logotipo Orange.....	31
Figura 12: Entrada principal ESPOCH	33
Figura 13: Genero.....	41
Figura 14: Edad	42
Figura 15: Ingresos mensuales	43
Figura 16: Lugar de conexión a Internet.....	44
Figura 17: Días de conexión a Internet.....	45
Figura 18: Conocimiento de Internet	45
Figura 19: Horas al día de Internet	46
Figura 20: Años de uso de Internet	46
Figura 21 :Horas de búsqueda de información.....	48
Figura 22 : Horas de chat académico.....	49
Figura 23 : Consultas académicas a compañeros.....	50
Figura 24 : Videos académicos	50
Figura 25: Horas de chat por entretenimiento	52
Figura 26: Horas de redes sociales	52
Figura 27 : Horas por semana de videos en Youtube	53
Figura 28 : Rendimiento académico	54

TERMINOLOGÍA

3D.- Tecnología de visualización en tres dimensiones.

ACJ.- Análisis Clúster Jerárquico.

B-LEARNING.- Aprendizaje virtual semipresencial.

CLUSTERING.- Técnica de agrupamiento por conglomerados comunes.

CONATEL.- Consejo Nacional de Telecomunicaciones.

DTIC.- Dirección de Tecnologías de la Información y Comunicación.

EDUREDES.- Redes sociales educativas.

E-LEARNING.- Educación virtual a distancia.

ESPOCH.- Escuela Superior Politécnica de Chimborazo.

EVA.- Entorno Virtual de Aprendizaje.

INEC.- Instituto de Estadística y Censos.

KDD.- Por sus siglas en inglés Knowledge Discovery from Databases.

K-MEDIAS.- Método de clustering.

LINKEDIN.- Red social profesional.

M-LEARNING.- Aprendizaje electrónico móvil.

OLAP.- Por sus siglas en inglés Online Analytical Processing.

SPSS.- Programa estadístico informático.

TCP/IP.- Protocolo de control de transmisión/Protocolo de Internet.

TIC.- Tecnologías de la Información y Comunicación.

UNESCO.- Organización de las Naciones Unidas.

UTPL.- Universidad Técnica Particular de Loja.

WWW.- World Wide Web

ÍNDICE DE ANEXOS

Anexo 1: Entrevista.....	81
Anexo 2: Encuesta.....	82
Anexo 3: Estadísticos descriptivos en SPSS.....	84
Anexo 4: K-Medias para el uso de Internet en actividades académicas	92
Anexo 5: K-Medias para el uso de Internet en actividades de entretenimiento.....	96
Anexo 6: Pruebas chi-cuadrado y medidas simétricas.....	100
Anexo 7: Regresión Binomial Hipótesis 1.....	102
Anexo 8: Regresión Binomial Hipótesis 2.....	105
Anexo 9: Regresión Binomial Hipótesis 3.....	108
Anexo 10: Regresión Binomial Hipótesis 4.....	111

RESUMEN

Ecuador es un país que se encuentra en vías de desarrollo y que está implementando políticas enfocadas en la reducción de la brecha digital, lo que ha dado paso a la inclusión del uso de la tecnología en todos los niveles de la educación, generando nuevos retos para docentes y estudiantes, quienes se ven en la obligación de inmiscuirse en nuevas metodologías para el aprendizaje, por ello en la Escuela Superior Politécnica de Chimborazo, se investiga acerca de la incidencia de los factores económicos y los usos del Internet sobre el rendimiento académico; para lo que se utiliza la metodología Knowledge Discovery in Databases (KDD) de la minería de datos que permite el descubrimiento del conocimiento, a partir de los datos obtenido de las encuestas aplicadas en forma aleatoria a los estudiantes, mediante la técnica de clustering se agrupo los datos de manera homogénea para crear perfiles que se contrastan con los ingresos económicos y el rendimiento académico, posteriormente a través de las pruebas de hipótesis se determina la incidencia entre las variables de estudio, obteniendo que: los ingresos económicos de las familias no incide en el uso del Internet para el aprendizaje, pero si incide en el uso del Internet para el entretenimiento y que el uso de la tecnología no incide en el rendimiento académico.

PALABRAS CLAVES: Internet, tecnología, minería de datos, k-means, clustering, rendimiento académico, educación, TICs.

ABSTRACT

Ecuador is a developing country and is implementing policies focused on reducing the digital divide, it has led to the inclusion in use of technology at all levels of education, it is creating new challenges for teachers and students who are obliged to know new methodologies for learning, so in the Escuela Superior Politécnica de Chimborazo, it has investigated about the impact of economic factors and uses of the Internet on academic performance; The KDD methodology of data mining allows to get knowledge from data obtained from surveys applied in randomly to students, using the technique of clustering data homogeneously, it has grouped in profiles and they are contrasted with the income and academic performance, later through hypothesis tests incidence among the study variables is determined: the income of families does not affect the use of the Internet for learning, but it affects the use of the Internet for entertainment and the use of technology does not affect academic performance.

KEYWORDS: Internet, technology, data mining, k-means, clustering, academic performance, TICs.

CAPITULO I
PLANTEAMIENTO DE LA INVESTIGACIÓN

1. Introducción

El presente documento está compuesto por cinco capítulos, distribuidos de forma organizada para permitir al lector una mejor comprensión sobre el presente trabajo de titulación denominado: “Análisis de los usos de la tecnología en la Escuela Superior Politécnica de Chimborazo (ESPOCH)”, en el que se analiza como inciden los factores económicos en el uso del Internet y como incide el uso de la tecnología en el rendimiento académicos de los estudiantes.

A continuación, se presenta el contenido de cada uno de los capítulos:

En el Capítulo I, se presenta la problemática que es el objeto de estudio del presente trabajo de investigación y sus alcances delimitados mediante el planteamiento de objetivos, sustentados por las preguntas de investigación con sus correspondientes hipótesis.

El Capítulo II, abarca la conceptualización teórica de los componentes a investigar, desde una aproximación a lo que es la llamada brecha digital, pasando por el uso de Internet, hasta sus aplicaciones en la educación, además se estudió como parte de la minería de datos la metodología conocida como el proceso de extracción del conocimiento, Knowledge Discovery in Databases (KDD).

En el Capítulo III, se define la metodología empleada para la investigación, en la que se determinó la población y la muestra para el estudio, también las herramientas para la recolección de la información, definición de variables y técnicas utilizadas para abstraer el conocimiento a partir de los datos seleccionados.

El Capítulo IV, desarrolla la etapa de análisis estadístico y minería de datos, utilizando herramientas como el software estadístico SPSS para aplicar métodos estadísticos como: conteo de frecuencias, análisis clúster, determinación del discriminante y la explotación del conocimiento.

En el Capítulo V, se discute los resultados obtenidos, llegando a plantear conclusiones y proponer trabajos futuros fruto de esta investigación.

El resultado del presente trabajo servirá para determinar, en que forma los factores socioeconómicos inciden en el rendimiento académico y en los usos que se le da al Internet.

1.1. Problemática

La sociedad está cambiando, es una de las frases perennes que durante años se ha mencionado tanto en temas sociales, culturales y tecnológicos; en las últimas décadas se ha producido una gran revolución tecnológica que está acompañada de la masificación del acceso al Internet, lo que conlleva a nuevas formas de aprendizaje, ya sea de contenido bueno o malo; acceso a redes sociales, donde se comparte todo tipo de información o se puede contactar a diversos grupos de personas; sitios de videos educativos o de entretenimiento; enlaces de video en tiempo real, donde se puede realizar una conferencia científica o ver un partido de fútbol, etc.

Dicha revolución se involucra en la vida de las personas hasta el punto de cambiar totalmente sus hábitos por la influencia de estas herramientas tecnológicas, las nuevas generaciones no desean perderse lo que pasa en todo el mundo, por lo que día a día viven más conectados a la web; la proliferación y el mal uso de las redes sociales desconectan al internauta totalmente su entorno, haciendo que los detalles de la vida pasen inadvertidos.

“¡Otro viernes por la mañana! Me levanto a la hora de costumbre, voy a la ducha, me pongo camisa y corbata, porque tengo clases, mientras enciendo la computadora preparo mi taza de café, reviso mis notificaciones de Facebook, el correo, el EVA de la UTPL, y vuelo hasta el punto de reunión virtual, donde me encuentro con mis compañeros, tutor y uno que otro ser extraño de otro mundo que le interesó la clase; activo el intercomunicador y escucho que discuten sobre los problemas que tuvieron los compañeros al instalar la plataforma (...). Antes de concluir la clase, el tutor nos recuerda que, para quienes no hayan podido asistir, podrán ver el video o leer el log de la conversación. Se despidió y finalizo la conexión. ¡Otra vez, se enfrió mi café!” (Garcés H., 2011).

En Internet, se puede hacer muchas cosas, el problema está en si lo que se hace es de provecho o no, pues se puede invertir tiempo en estudiar, prepararse para un examen o ver videos de una serie que aún no se publica en televisión abierta, pero, ¿qué determina este comportamiento? ¿influyen los recursos económicos del internauta?

Si se usa el Internet de forma adecuada con un enfoque académico, ¿mejora el rendimiento de los estudiantes? En la presente investigación, se buscó dar respuesta a algunas de estas interrogantes.

1.2. Objetivos

1.2.1. General

- Determinar los usos de Internet en los estudiantes de la Escuela Superior Politécnica de Chimborazo.

1.2.2. Específicos

- Determinar la incidencia de la variable de los ingresos económicos, sobre los usos del Internet para el aprendizaje.
- Determinar la incidencia de la variable de los ingresos económicos, sobre los usos del Internet para el entretenimiento.
- Determinar la incidencia del uso de la tecnología en el aprendizaje, sobre el rendimiento académico.
- Determinar la incidencia del uso de la tecnología para el entretenimiento, sobre el rendimiento académico.

1.3. Preguntas de investigación

- ¿Cómo se relacionan los niveles de ingreso de las familias de los estudiantes universitarios, con los usos de Internet en actividades académicas y de entretenimiento?
- ¿Cómo se relacionan el rendimiento académico, con los usos de Internet en actividades académicas y de entretenimiento?

1.4. Hipótesis

- Hipótesis 1: El nivel de ingresos económicos determina como se utiliza Internet para el aprendizaje.
- Hipótesis 2: El nivel de ingresos económicos determina como se utiliza Internet para entretenimiento.
- Hipótesis 3: El uso de la tecnología en el aprendizaje incide en el rendimiento académico.
- Hipótesis 4: El uso de la tecnología para entretenimiento incide en el rendimiento académico.

CAPITULO II
MARCO TEÓRICO Y CONCEPTUAL

2. Marco teórico y conceptual

2.1. Brecha digital

La brecha digital se define como la separación que existe entre las personas (comunidades, estados, países...) que utilizan las Tecnologías de la Información y Comunicación (TIC) como una parte rutinaria de su vida diaria y aquéllas que no tienen acceso a las mismas y que, aunque las tengan no saben cómo utilizarlas (Castaño, & others, 2008).

Por lo tanto, se puede decir brecha digital a la desigualdad entre las personas que pueden tener acceso o conocimiento en relación a las nuevas tecnologías y las que no, dicha distinción puede ser determinada por la realidad socioeconómica de la gente, así como al nivel de capacidad tecnológica y alfabetización estos parámetros pueden determinar el acceso a ordenadores personales, la telefonía móvil, la banda ancha y otros dispositivos.

El crecimiento e inserción de las TIC en la economía mundial, han generado condiciones que afectan profundamente a nuestra sociedad, dividiéndola entre comunidades que apropian efectivamente estos recursos y aquellos que no lo hacen, situación denominada brecha digital (Zapata, & Rojas, 2014).

La brecha digital se reduce a medida en que la gran mayoría de la población tenga un acceso básico o mínimo al entorno digital; sin embargo, quedan desigualdades en la capacidad de acceso que pueden marcar diferencias decisivas en las oportunidades sociales, laborales y culturales de las personas. De aquí la importancia de reducir o eliminar la brecha digital lo cual será solo posible si la sociedad y el gobierno trabajan conjuntamente para hacer que la tecnología esté al servicio y engrandecimiento de la sociedad.

2.2. Tipos de brecha digital

Se proponen tres tipos de brecha digital: la de acceso, basada en la diferencia entre las personas que pueden acceder y las que no a las TIC; la de uso, basada en las personas que saben utilizarlas y las que no ; y las de la calidad del uso, basada en las diferencias entre los mismos usuarios (Camacho, 2005).

Brecha digital de acceso

Es la posibilidad o dificultad de que la población disponga de herramientas tecnológicas actuales, lo que depende de una combinación de factores económicos, el limitante de una

buena infraestructura de telecomunicaciones e informática, así como la diferencia en el nivel de educación.

El acceso ha marcado dos grupos diferenciados que son los incluidos y los excluidos a las nuevas tecnologías como son los dispositivos computadores personales, smartphones y acceso al Internet.

Brecha digital de uso

Se la puede describir como el conocimiento y práctica de la persona para utilizar la tecnología y acceder a la información que necesita. Aún existe una cantidad considerable de personas que carecen de accesibilidad al Internet, debido a la ubicación, conocimiento o el nivel socioeconómico.

Brecha digital de calidad de uso

A pesar que gran parte de la población tiene acceso a la información y sabe cómo llegar a ella, el uso en muchos casos de esta información no es el adecuado, o permite que se genere un desperdicio de este recurso tan valioso pues no siempre se lo aprovecha ni se explota su potencialidad.

2.3. Factores de la brecha digital

Existen varios factores que inciden en la brecha digital, entre los más importantes se destacan:

Recursos económicos

Posibilitan que tengan o no acceso a las tecnologías, debido a que el nivel económico que presente el individuo determina el poder adquirir equipos informáticos de vanguardia, de mediana tecnología o no poder adquirirlos.

La economía de un hogar permite o no pagar por el servicio de acceso a la red y determina la posibilidad de asistir a capacitaciones en tecnología que tengan un valor económico.

Geografía

Los residentes que tiene mayor acceso son los de los países desarrollados en comparación a aquellos que habitan en países en vías de desarrollo; y ya dentro de los países existen zonas generalmente las rurales sin la infraestructura necesaria para establecer acceso a Internet, además la forma de vida y las actividades económicas de la que el sector depende como suelen ser las agrarias o la ganadería, son factores que establecen una dificultad más grande

para que sus moradores conozcan acerca de las TICs.

Edad

Las personas de más edad tienen menos posibilidades debido a que en su mayoría nunca han tenido el acceso a las TICs y de tener acceso a ellas, existe el limitante de la falta de conocimiento para usarlos y en otros casos simplemente muestran desinterés por aprender. Al contrario de las nuevas generaciones que cada vez a más temprana edad empiezan a acceder y manipular las nuevas tecnologías.

Nivel educativo

El grado de aprovechamiento de las TICs depende de la alfabetización del individuo en muchos casos la falta de pilares básicos de la educación no permite apreciar y manejar la información de la red.

Ancho de banda

El ancho de banda se define como la capacidad en Hertz necesaria para acceso a Internet, está relacionado a la velocidad y eficiencia con que se accede y se trabaja en Internet. Cada vez el tráfico de datos sobrepasa el tráfico de voz, el 2000 fue el año histórico en que sucedió por primera vez esto entre el tráfico de datos y voz (Santoyo, & Martínez, 2003).

Género

La ciencia y la tecnología son áreas de conocimiento que tradicionalmente han estado vinculadas a lo masculino. Las cifras estadísticas de científicas y tecnólogas corroboran esta escasa presencia de mujeres en este tipo de profesiones y esconden aún más los pocos ejemplos a seguir que existen de mujeres científicas e ingenieras.

Por tradición la mujer al ser responsable de las tareas caseras ha provocado la falta de tiempo y desinterés por aprender nuevas tecnologías y ha sido un fenómeno que incluso hasta la actualidad ha permanecido. Por otro lado, el género femenino ha sido discriminado ante el cambio en las empresas y las condiciones salariales.

Estamos inmersos en una sociedad de la información en la que hay muchos grupos excluidos, pero no se trata solo de un problema de falta de recursos, infraestructuras o conocimientos. Se trata también de un problema social. Las nuevas tecnologías aparecen en una sociedad ya desigual y, por tanto, generan resultados desiguales (Castaño, & others, 2008).

2.4. Internet

Internet es un conjunto descentralizado de redes de comunicación interconectadas, que utilizan la familia de protocolos TCP/IP, garantizando que las redes físicas heterogéneas que la componen funcionen como una red lógica única, de alcance mundial. Uno de los servicios que más éxito ha tenido en Internet ha sido la World Wide Web (WWW). La WWW es un conjunto de protocolos que permite, de forma sencilla, la consulta remota de archivos de hipertexto («La internet...», 2009).

Con la Internet se puede eliminar las jerarquías, las diferencias y las distancias, ya que favorece la libertad de expresión y permite que la información, el conocimiento y el mundo entero estén al alcance de un mayor número de personas, es un instrumento que se puede ejecutar desde casi cualquier parte del mundo, ya sea por computadoras, como por dispositivos móviles como son tablets, teléfonos inteligentes.

Uso del Internet

Internet, entonces, no es una herramienta tecnológica más, ahora es un medio de comunicación, interacción y organización social; razón por la cual hace que su popularidad siga incrementando y más parte de la población busque acceder a esta red, pues Internet es un espacio en el cual se reflejan los distintos ámbitos de la vida humana: la política, la educación, el esparcimiento, las cosas buenas y las cosas malas. Por lo que se le puede dar una infinidad de usos entre las que podemos nombrar:

Estudio

El estudio ha sido revolucionado con el uso del Internet debido a que permite realizar consultas, visitar foros educativos, así como facilita la educación online es decir hoy en día es posible recibir clases que están siendo dictadas a distancias considerables o incluso en otros países, por medio de video conferencias.

Trabajo

Hoy en día no hay empresa que se abstenga del uso de Internet, en las empresas, agiliza trámites y transferencias de información, mejorando así a la eficiencia empresarial.

Comunicación

La comunicación es la base de las relaciones humanas y mediante el uso del Internet las vías de comunicación han aumentado considerablemente por la facilidad de mantener y relacionar personas por medio de video llamadas, chats, redes sociales, entre otras.

Ocio

Internet ofrece la oportunidad de almacenar en un mismo lugar toda la información de diferentes tópicos como música, arte, cultura, deportes medicina, literatura, política, ingeniería y mucho más. Por medio de texto audio video, música imágenes, juegos en línea. Por lo cual el usuario usa el Internet para informarse, entretenerse y divertirse.

La gente debe tener conciencia que en el Internet se puede exponer a peligros para los cuales se hace necesario tener un comportamiento adecuado y responsable hacia lo que vemos, leemos y compartimos. La sinceridad, el respeto y la tolerancia son valores importantes a tener y promover cuando se usa la red.

2.5. Uso del Internet en el Ecuador

En el Ecuador está estipulado en el Reglamento General de la ley Especial de Telecomunicaciones que dispone el Consejo Nacional de Telecomunicaciones (CONATEL) dicta la regulación de acceso al servicio de Internet, y que la interconexión y conexión se permitirán en condiciones de igualdad, no discriminación, neutralidad y libre y leal competencia, a cambio de la debida retribución (Soline Moreno, 2016).

En Ecuador un 46,35% de personas tuvieron acceso a Internet, de acuerdo a datos del Instituto de Estadísticas y Censos (INEC). El crecimiento de esta cifra se debe principalmente al incremento de hogares con acceso a Internet, así como a las políticas públicas del gobierno en proveer de Internet a lugares públicos.

El INEC también arroja el dato que en el Ecuador se redujo el analfabetismo digital de 32,42% en 2008 a 14,43% en 2014 (INEC, 2016)

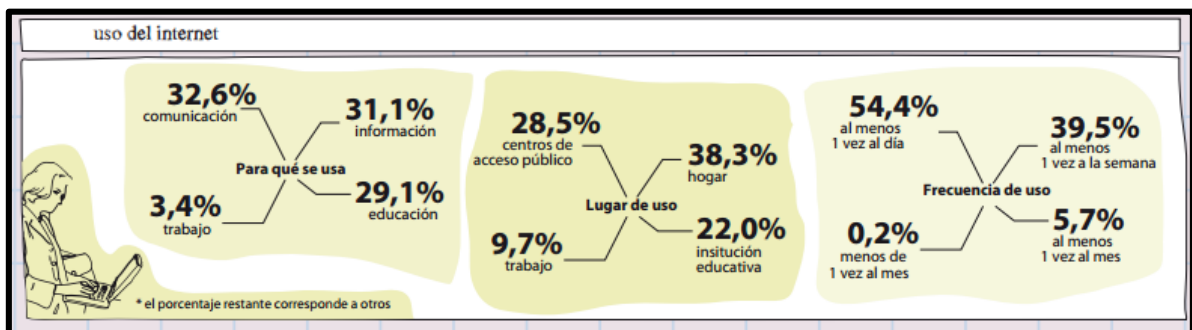


Figura 1: Uso del Internet en el Ecuador
Fuente: (INEC, 2014)

2.6. Tecnología y educación

La sociedad actual, demanda cambios en los sistemas educativos de forma que estos se sean más flexibles y accesibles, menos costosos que se puedan acceder con facilidad en cualquier momento. Las instituciones de formación educativa deben responder a estos desafíos que se presentan, promoviendo nuevas estrategias y metodologías apoyadas en las tecnologías de información y comunicación (TICs).

El encuentro entre educación y nuevas tecnologías genera una poderosa industria; la industria educacional. La convergencia entre dos sectores que combinadamente gastan alrededor del diez por ciento del producto interno de los países, generando en su entorno, una serie de transformaciones y el aura de una modernidad global (Brunner, Brunner, & Tedesco, 2003).

Las TICs en la educación presentan grandes ventajas como poder simplificar actividades y ampliar la capacidad del individuo; ayudar para que los estudiantes comprendan mejor el objeto de estudio y puedan aplicarlo, entregan la opción de organización para llevar la materia, organizarla y simplificarla. Las TICs permiten que se obtenga información de fuentes directas en tiempos mínimos y como lo menciona Medina en su blog dice que las tecnologías no cambian la manera de enseñar, pero pueden ayudar a introducir nuevas metodologías y entornos de aprendizaje (Medina, 2011).

2.7. Áreas de desarrollo prioritario para la integración de las nuevas tecnologías en la educación.

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) determina que las tecnologías de la información y la comunicación (TIC) pueden contribuir al acceso universal a la educación, la igualdad en la instrucción, el ejercicio de la enseñanza y el aprendizaje de calidad y el desarrollo profesional de los docentes, así como a la gestión dirección y administración más eficientes del sistema educativo (UNESCO, 2016). Para este propósito ha determinado dos áreas de desarrollo prioritario las nuevas prácticas educativas y la medición de aprendizajes.

2.8. Nuevas prácticas educativas

Es importante considerar la manera en que las TICs pueden favorecer el desarrollo de nuevas prácticas educativas, haciéndolas más pertinentes y eficaces, para esto se debe fortalecer el protagonismo que tienen los docentes en los cambios educativos. Los maestros son quienes analizan los factores que intervienen en su desempeño, lo que afecta al rendimiento de sus estudiantes, y pueden reconocer, multiplicar y potenciar aquellas experiencias de aprendizaje

que las TICs posibilitan mejorar con respecto a las prácticas tradicionales de enseñanza (Severin, 2013).

Las políticas educativas, así como el ingenio del catedrático al impartir el conocimiento en los alumnos utilizando herramientas como presentaciones multimedia, videos, música, imágenes, simuladores de laboratorio, aulas virtuales etcétera, van creando y revolucionando el sistema de enseñanza y aprendizaje. Es decir se necesita de profesionales, que sean capaces de producir cambios, para ello hay que sensibilizarlos en la búsqueda de soluciones a diversos problemas.

2.9. Evaluación de aprendizajes

La evaluación es una actividad sistemática y continua que tiene como misión recoger información fidedigna sobre el proceso de aprendizaje en su conjunto para ayudar a mejorar el propio proceso, y dentro de él, los programas, las técnicas de aprendizaje, los recursos, los métodos y todos los elementos del proceso (Martínez & Salanova, s. f.).

Es fundamental indagar qué resultados se obtienen con el uso de TICs tanto para los docentes y los estudiantes, esta información, es analizada para la toma de decisiones que influirán en el proceso educativo, como, por ejemplo, en la planeación de los contenidos a desarrollar, el refuerzo de temas, el uso de nuevas técnicas o enfoques, o bien para seguir en la marcha, también permite la obtención de información acerca de los intereses y necesidades de los estudiantes.

Las TICs ofrecen oportunidades innovadoras para la evaluación del aprendizaje de cada estudiante y del desempeño de los docentes, las escuelas y los sistemas educativos. Es posible realizar pruebas nacionales e internacionales destinadas a medir los progresos en el aprendizaje mediante la incorporación de nuevas formas de aplicación, menor costo, mayor periodicidad, sino por herramientas más poderosas de análisis de los resultados y por tanto de apoyo a la gestión de los aprendizajes (Severin, 2013).

2.10. Educación virtual

La educación virtual es un medio pedagógico donde los contenidos se presentan a través de las redes de comunicación, influenciada principalmente por Internet, donde la comunicación es mediada por uno o más ordenadores y el entorno de enseñanza y aprendizaje se realiza en una aula virtual.

2.10.1. E-learning.

E-learning, es una modalidad de enseñanza-aprendizaje que consiste en el diseño, puesta en práctica y evaluación de un curso o plan formativo desarrollado a través de redes de ordenadores y puede definirse como una educación o formación ofrecida a individuos que están geográficamente dispersos o separados o que interactúan en tiempos diferidos del docente empleando los recursos informáticos y de telecomunicaciones. Lo característico del e-learning es que el proceso formativo tiene lugar totalmente o en parte a través de una especie de aula o entorno virtual en el cual tiene lugar la interacción profesor-alumnos, así como las actividades de los estudiantes con los materiales de aprendizaje (Moreira & Segura, 2009).

Ventajas

- Permite reducir y hasta eliminar gastos de traslado, alojamiento, material didáctico.
- Combina el poder de Internet con el de las herramientas tecnológicas.
- Los usuarios pueden acceder al contenido desde cualquier conexión a Internet, cuando les surge la necesidad.
- Anula las distancias geográficas y temporales.
- Permite utilizar la plataforma con mínimos conocimientos.
- Ofrece libertad en cuanto al tiempo y ritmo de aprendizaje.

Desventajas

- Mayor trabajo de planificación y desarrollo del curso.
- Necesidad de un equipo técnico de producción y de gestión.
- Mayor esfuerzo de los docentes, ya que deben ser expertos en su materia y el uso de las TICs.
- Mayor esfuerzo de los alumnos al adaptarse la metodología.
- Requerimiento de estudiantes organizados y autónomos.
- El alumno necesita estar motivado para aprender de otra manera puede optar por la renuncia.

2.10.2. B-learning.

B-learning o blended-learning es la combinación del aprendizaje presencial con el aprendizaje a distancia utilizando en esta la red Internet (Moreira, & Segura, 2009).

La idea es utilizar lo mejor de la formación a distancia y lo mejor de la formación presencial, para de esta manera lograr un método de educación muy adaptable y completo. El uso de

las TIC se utiliza durante todo el proceso de formación b-learning, no únicamente en los procesos de formación a distancia.

Ventajas

- Reducción de costos.
- Optimización del uso y la programación del tiempo.
- Reducción de sesiones presenciales.
- Comunicación y asesoría permanentes con el docente.

Desventajas

- Se necesita tiempo para que los alumnos se adapten a esta forma de estudio.
- Diferencia en el conocimiento de utilización de las TICs.
- Dependencia de la infraestructura computacional y a la disponibilidad de acceso a Internet de los participantes.
- Falta de participación y compromiso en las asignaturas.
- Los alumnos no están acostumbrados al autoaprendizaje.
- Los educadores no están bien preparados en el manejo de las TICs.
- La evaluación que el docente ha de hacer sobre sus alumnos, implica un incremento del tiempo que el educador destina a la evaluación.

2.10.3. M-learning.

Mobile Learning es el concepto utilizado para referirse a los ambientes de aprendizaje basados en la tecnología móvil, enfocados a impulsar y mejorar los procesos de aprendizaje.

La mayor parte de los proyectos se han centrado en un modelo instruccional de transmisión de información, donde el profesor produce contenidos y los hace accesibles a los alumnos para su consumo a través del dispositivo. Esto ha llevado a que los principales usos de este tipo de herramientas se hayan limitado a la consulta de datos, la organización administrativa y a la interacción guiada a través de la respuesta a cuestionarios (Izarra, 2010).

El M-Learning es una realidad que puede ampliar el tiempo potencial del que disponen los estudiantes. Además, da la posibilidad de usarlo en ambientes integrados de e-learning o como un ambiente aislado de aprendizaje.

Para un correcto empleo de la metodología M-learning los docentes deben estar claros que la información que deben compartir debe ser simple debido a que el acceso en dispositivos

móviles se realiza en momentos breves y con una pantalla reducida, por lo que conviene dar prioridad a lo importante.

Tener claras las limitaciones y oportunidades del uso del M-Learning entrega la posibilidad de crear estrategias que realmente brinden un valor agregado a los estudiantes y profesores en sus procesos de enseñanza y aprendizaje.

Ventajas

- El teléfono móvil está disponible las 24 horas del día.
- La variedad de juegos generados para móviles, impulsa la creatividad y la colaboración.
- Actividades online de formación están disponibles para dispositivos móviles.
- Acceso inmediato a mensajes, correos, recordatorios y noticias generados en tiempo real.
- Acceso a datos en línea para apoyar el trabajo de campo.
- Contacto inmediato con los padres y representantes.
- Participar en escenarios educativos mediante el uso del chat.

Desventajas

- Pantallas pequeñas. Esto conlleva dificultades en la lectura de textos medianos, la cantidad de información visible es limitada y el desplazamiento continuo por la pantalla para leer toda la información.
- Los dispositivos móviles pueden convertirse en una fuente de distracción.
- Actualmente las baterías son poco duraderas de los dispositivos móviles.
- Dificultades o imposibilidad de instalar y usar determinado software.

2.11. Las redes sociales en la educación

Las redes sociales son estructuras sociales web que favorecen el contacto entre personas por diferentes motivos o intereses como pueden ser amistad, parentesco o que compartan el mismo conocimiento. Las redes sociales están introduciendo nuevas formas de entender las relaciones personales y la comunicación social.

Existen redes sociales, dedicadas a la cultura, como Linkara, o a las relaciones laborales, como LinkedIn y otras son las educativas en este marco destacamos EDUREDES. En la educación las redes sociales permiten el incremento de la eficacia del uso práctico de las TIC.

A pesar de todas las ventajas que las redes sociales tienen en la educación, existe un porcentaje considerable de docentes que sienten temor al utilizarlas por desconocimiento de

las posibilidades que ofrecen y no saber cómo utilizarlas. Mientras que para las nuevas generaciones de estudiantes el uso de estas redes sociales es algo tan básico y sienten necesidad que el uso de ellas este dentro de su formación educativa.

En esta etapa de cambio en el sistema de educación que vivimos dentro del Ecuador, se necesita la colaboración de todo el cuerpo educativo tanto en hábitos como en actitud, para así incorporar y explotar su potencial.

2.12. Mundos virtuales en la educación

Un mundo virtual es un entorno simulado, es una representación en 3D, accesible por ordenador, en el que los participantes participan interactivamente en tiempo real con el uso de un avatar, los mundos virtuales enfocados en la educación serían los que supervisan el uso del avatar para cumplir actividades de aprendizaje o misiones que busquen el cumplimiento de objetivos que sirvan para la formación académica (López, 2011).

Dentro de los mundos virtuales más utilizados para la educación esta OpenSim y SecondLife, donde podemos encontrar bibliotecas virtuales, sitios arqueológicos, salas de colaboración, aulas virtuales y material compartido entre los internautas.

En estos mundos se puede desde dictar una clase, y por medio de los avatares estar sentados atendiendo a la clase en un salón virtual semejante a la educación tradicional solo que desde la propia casa u oficina de trabajo.

2.13. Aproximación al concepto de minería de datos

La minería de datos puede definirse inicialmente como un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias al examinar grandes cantidades de datos. La disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas ha transformado el análisis de datos orientándolo hacia determinadas técnicas especializadas englobadas bajo el nombre de minería de datos o data mining (Pérez López, & Santín Gonzáles, 2007).

La minería de datos ayuda a reducir de manera automática o semiautomática todos estos datos en información útil que previamente fue desconocida y esta información se puede utilizar para entender, buscar soluciones y sacar conclusiones a determinado problema.

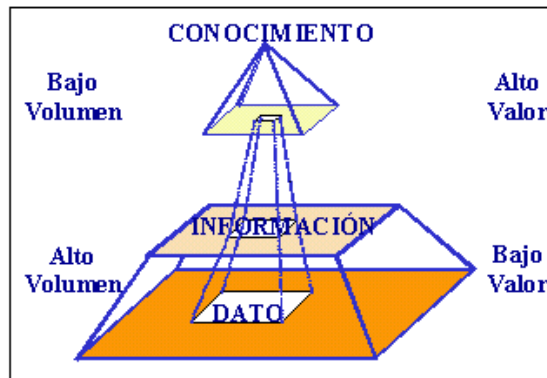


Figura 2: Relación entre dato, información y conocimiento
Fuente: (Molina, 2002)

La minería de datos, desde un punto de vista estrecho, comprende, como sistema de extracción de relaciones, los métodos basados en la computadora, requiriendo poca involucración y ayuda por parte del analista en la obtención de información relevante. Incluiríamos los algoritmos de neuronas artificiales, arboles de decisión, inducción de reglas, lógica difusa, Análisis de algoritmos genéticos (Fernández, 2003).

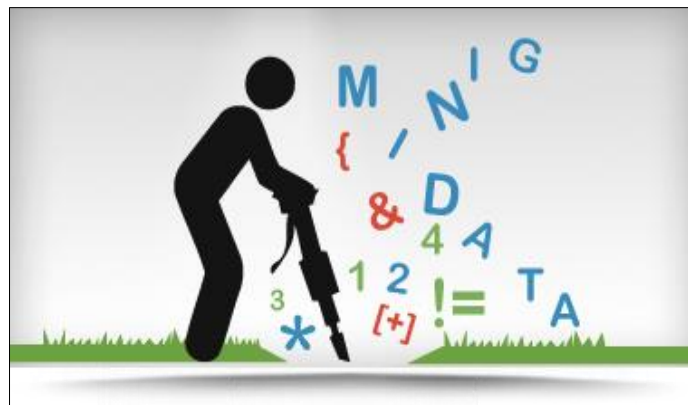


Figura 3: Minería de Datos
Fuente: ("Minería de Datos", s/f)

2.14. Proceso de extracción del conocimiento a partir de datos (KDD)

La extracción de la base del conocimiento KDD se refiere al proceso no—trivial de descubrir conocimiento e información potencialmente útil dentro de los datos contenidos en algún repositorio de información. No es un proceso automático, es un proceso interactivo que exhaustivamente explora volúmenes muy grandes de datos para determinar relaciones. Es un proceso que extrae información de calidad que puede usarse para dibujar conclusiones basadas en relaciones o modelos dentro de los datos («KDD Proceso de Extracción de conocimiento | WebMining», 2011).

Este proceso consta de varias fases como la preparación de datos, limpieza de los datos, integración de datos, selección de los datos, transformación de los datos, minería de datos, evaluación de patrones y presentación del conocimiento. Además, el proceso de extracción del conocimiento incorpora diferentes técnicas como árboles de decisión, regresión lineal, redes neuronales artificiales, técnicas bayesianas, máquinas de soporte vectorial, etc., de campos diversos (aprendizaje automático e inteligencia artificial), estadística, bases de datos, etc., y aborda una tipología variada de problemas, categorización, estimación, regresión, agrupamiento (Pérez López, & Santín Gonzáles, 2007).

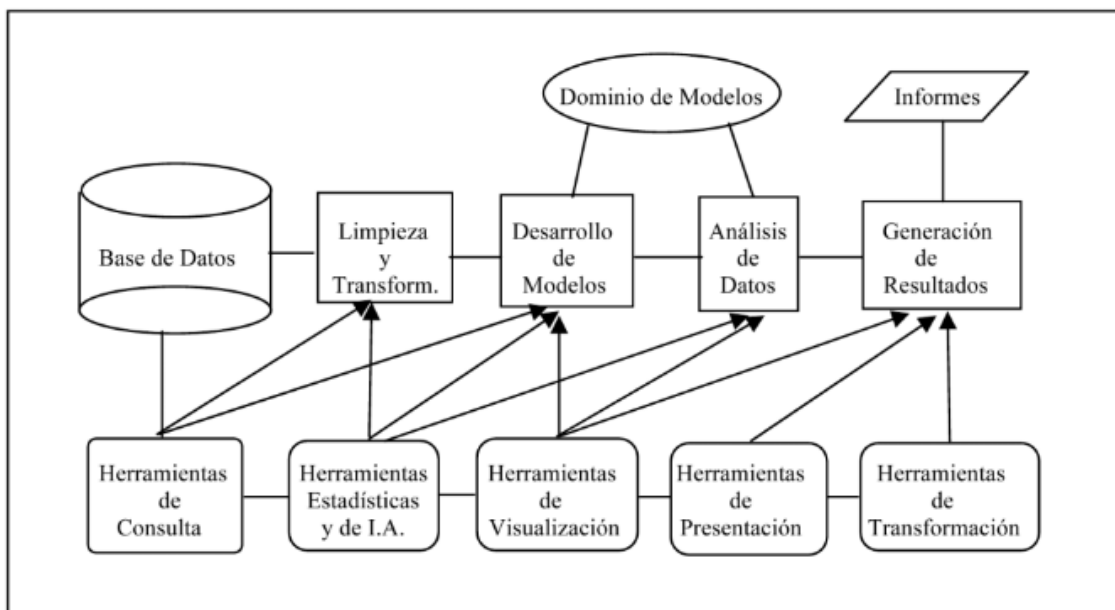


Figura 4: Etapas del KDD
Fuente: (Pérez López & Santín Gonzáles, 2007)

2.14.1. Selección de los datos.

Esta fase determina que las fases sucesivas sean capaces de extraer conocimiento válido a partir de la información original. Para la selección se determina las fuentes de datos y el tipo de información a utilizar, es la etapa donde los datos relevantes para el análisis son extraídos. La disponibilidad de grandes volúmenes de información en esta fase lleva a la necesidad de usar técnicas de muestreo para la selección de datos (Pérez López & Santín Gonzáles, 2007).

2.14.2. Limpieza de los datos.

En esta fase el objetivo es detectar y tratar la presencia de valores atípicos, imputar la información faltante o valores perdidos, eliminar datos erróneos o irrelevantes. En esta fase se utilizan herramientas de consulta y herramientas estadísticas (Pérez López, & Santín Gonzáles, 2007a).

2.14.3. Transformación de los datos.

Consiste en el tratamiento preliminar de los datos, transformación y generación de nuevas variables a partir de las ya existentes con una estructura de datos apropiada. Aquí se realizan operaciones de agregación o normalización, consolidando los datos de una forma necesaria para la fase siguiente («KDD Proceso de Extracción de conocimiento | WebMining», 2011).

2.14.4. Minería de datos.

La minería de datos se lleva a cabo a partir del desarrollo de modelos predictivos y descriptivos y mediante el análisis de datos. Una vez recogidos los datos de interés, un explorador puede decidir qué tipo de patrón quiere descubrir, el tipo de conocimiento que se desea extraer va a marcar claramente la técnica de minería de datos a utilizar (Pérez López, & Santín Gonzáles, 2007a).

2.14.5. Evaluación de patrones.

Para seleccionar y validar los modelos extraídos de la fase de minería de datos es necesaria la fase consistente en el uso de criterios y evaluación de hipótesis. El modelo del despliegue a veces es trivial, pero otras veces requiere un proceso de interpretación o evaluación. En esta fase se utilizan adicionalmente herramientas estadísticas y de visualización (Pérez López, & Santín Gonzáles, 2007a).

2.14.6. Presentación del conocimiento.

La presente fase es la relativa a la difusión y uso del conocimiento que se deriva de las técnicas de minería de datos a través de los modelos correspondientes que habitualmente desembocan en la generación de resultados. El modelo puede tener muchos usuarios y necesitar difusión, con lo que puede requerir ser expresado de una manera comprensible para ser distribuido en la organización. En esta fase se utilizan herramientas de visualización, presentación y transformación de datos (Pérez López, & Santín Gonzáles, 2007a).

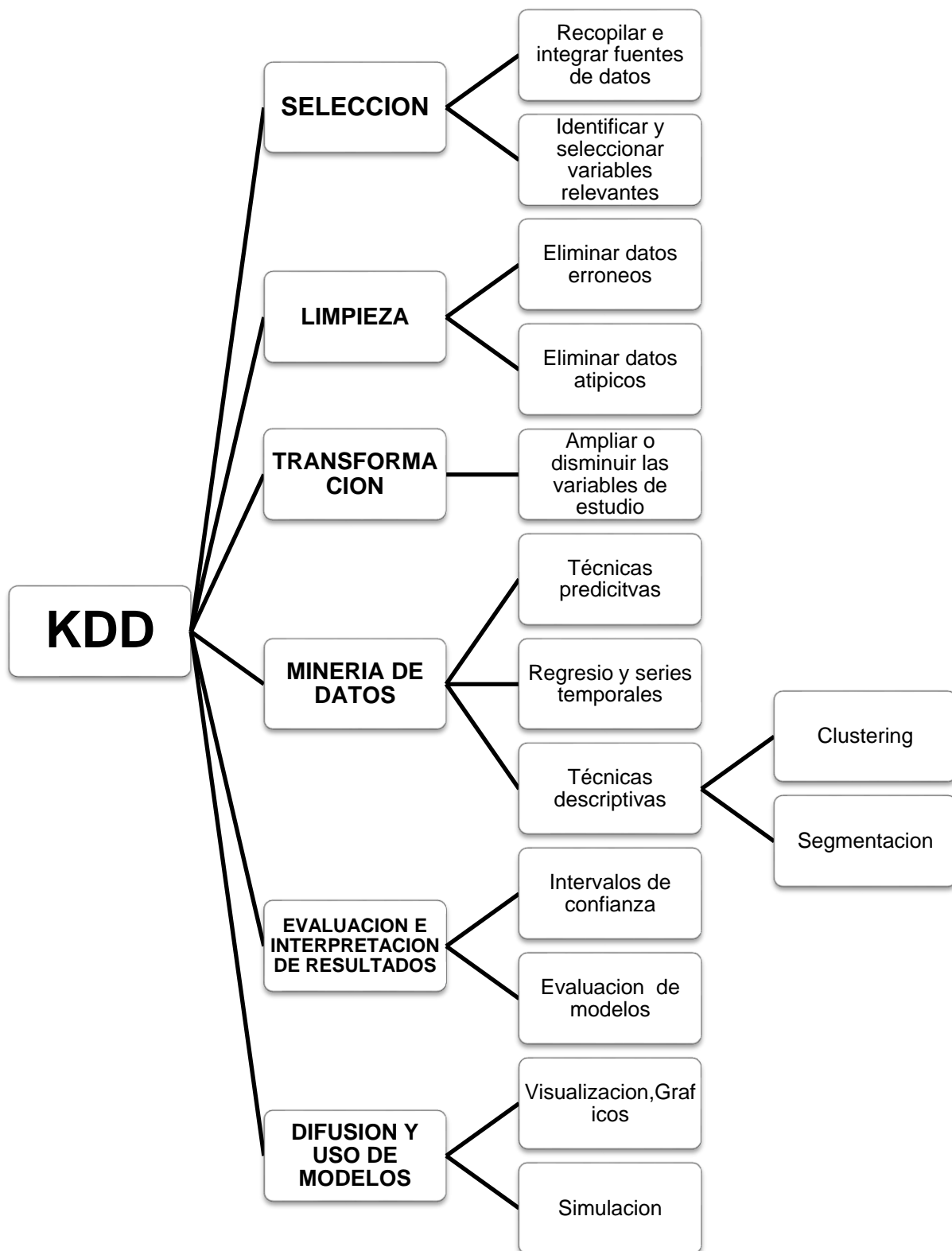


Figura 5: Resumen fases KDD
 Fuente: Elaboración Propia

2.15. Fuentes de datos

La minería de datos se aplica a todo tipo de datos imaginable: desde datos numéricos a imágenes de satélite, mamografías, música, archivos de ordenador, imágenes, etc. Podemos decir que cualquier cosa constituye un dato (Krall, 2006). Sin embargo, los datos por sí mismo no aportan conocimiento es necesario procesarlos y transformarlos.

Por lo cual el primer paso es almacenarlos para que lleguen a constituir una base de datos que aglomere el conjunto sin redundancias innecesarias en un soporte informático y accesible simultáneamente por distintos usuarios y aplicaciones. En una base de datos, los datos deben estar estructurados y almacenados de forma totalmente independiente de la aplicación que la utilizan (Cobo, s. f.).

Algunos de los tipos de bases de datos son:

Base de datos relacionales

Una base de datos relacional es una base de datos en donde todos los datos visibles al usuario están organizados estrictamente como tablas y en donde todas las operaciones que se realizan en la base de datos se hacen sobre estas tablas (Gonzales M., 2015).

Base de datos espaciales

Consiste en un conjunto de datos geográficos, estos datos son almacenados en cierto tipo de archivo que interpretan las aplicaciones geográficas que se encuentran en el mercado.

Bases de datos orientadas a objetos

Una base de datos orientada a objetos es una base de datos donde los elementos son objetos. Estos pueden ser bases de datos multimedia y contener videos, imágenes y sonidos, donde la herencia permite una mejor representación de la información. Estas bases de datos tienen identidad de ser un todo y no solo parte de una gran base, su ventaja frente a las bases de datos tradicionales es poder representar los modelos de datos con un marco mucho más eficiente, manteniendo la integridad y relación entre ellos (Gonzales M., 2015).

Base de datos documentales

Estas bases de datos permiten la indexación a texto completo y permiten realizar búsquedas más potentes, por ejemplo:

World Wide Web

Es un sistema de obtención y presentación de documentos distribuidos por Internet. Todos ellos forman un ingente entramado de información (Sánchez, s. f.).

2.16. Tipos de modelos de datos

Digamos que un modelo es una forma de aplicar un tratamiento a una cantidad masiva de datos para extraer información de ellos. Las finalidades de un modelo en minería de datos inciden en dar soporte al descubrimiento de relacionamientos, decisión, predicción y mejor de los procesos. La sistemática de cómo se llega a un modelo sigue el método científico: Datos, hipótesis y prueba de la hipótesis (Viera, Ortiz, & Ramirez, 2009).

2.16.1. Modelo descriptivo y sus operaciones.

Provee información sobre el estado actual del sistema.

Clustering o agrupamiento

El proceso de Clustering consiste en la división de los datos en grupos de objetos similares. Para medir la similaridad entre objetos se suelen utilizar diferentes formas de distancia: distancia Euclídea, de Manhattan, de Mahalanobis, etc. El representar los datos por una serie de clusters, conlleva la pérdida de detalles, pero consigue la simplificación de los mismos (Garre, Cuadrado, & Silicia, s. f.) .

Correlación

Los estudios correlacionales pretenden medir el grado de relación y la manera cómo interactúan dos o más variables entre sí. Estas relaciones se establecen dentro de un mismo contexto, y a partir de los mismos sujetos en la mayoría de los casos.

Asociación

Incluye técnicas conocidas como linkage analysis, utilizadas para buscar patrones que tienen una probabilidad alta de repetición.

2.16.2. Modelo predictivo y sus operaciones.

Provee información sobre estados futuros del sistema a partir de otras variables que cumplen el papel de predictoras. En este tipo de modelos es posible realizar las siguientes operaciones:

Regresión

Los métodos de análisis de series de tiempo son usados para relacionar los eventos con el tiempo.

Clasificación

Incluye los procesos de minería de datos que buscan reglas para definir si un ítem o un evento pertenecen a una clase de datos. Esta técnica, probablemente la más utilizada, incluye dos subprocesos: la construcción de un modelo y la predicción.

2.17. Técnicas de minería de datos

Las técnicas minería de datos se clasifican en técnicas predictivas, en las que las variables pueden clasificarse inicialmente en dependientes o independientes, técnicas descriptivas en las que todas las variables tiene inicialmente le mismo estatus y técnicas auxiliares (Pérez López, & Santín Gonzáles, 2007a).

2.17.1. Técnicas predictivas.

Las técnicas predictivas especifican los modelos para los datos en base a un conocimiento teórico previo. El modelo supuesto para los datos debe contratarse después del proceso de minería de datos antes de aceptarlo como válido (Pérez López, & Santín Gonzáles, 2007a).

Regresión

Es una técnica estadística utilizada para analizar la relación entre una variable dependiente o endógena métrica y varias variables independientes o exógenas también métricas. El objetivo esencia del análisis de la regresión múltiple es utilizar las variables independientes, cuyos valores son conocidos, para predecir la única variable criterio dependiente seleccionada por el investigador (Pérez López, & Santín Gonzáles, 2007a).

Series temporales

Una serie de tiempo corresponde a un conjunto de observaciones hechas respecto a una variable en momentos equidistantes en el tiempo.

Análisis de la varianza y covarianza

El análisis de la covarianza es una técnica estadística que, utilizando un modelo de regresión lineal múltiple, busca comparar los resultados obtenidos en diferentes grupos de una variable cuantitativa, pero corrigiendo las posibles diferencias existentes entre los grupos en otras variables que pudieran afectar también al resultado, las covariantes (Moliner, 2002).

Métodos bayesianos

Es un modelo gráfico que usa directamente los arcos para formar una gráfica a cíclica. Se usa muy frecuentemente las redes de Bayesiana cuando la incertidumbre se asocia con un resultado que puede expresarse en términos de una probabilidad (Molinero, 2002).

Algoritmos genéticos

Técnica computacional cuya utilización habitual se ha circunscrito al ámbito de la optimización, proporcionando soluciones mejoradas, no necesariamente óptimas.

Análisis discriminante

Es una técnica estadística utilizada para analizar la relación entre una variable dependiente o endógena no métrica (categórica) y varias variables independientes (o exógenas) métricas. El objetivo esencia del análisis discriminante es utilizar los valores conocidos de las variables independientes para predecir en que categoría de la variable dependiente se corresponden (Pérez López, & Santín Gonzáles, 2007a).

Arboles de decisión

Constituyen un método de segmentación, que trata de resolver los problemas de discriminación en una población segmentando de forma progresiva la muestra para obtener finalmente una clasificación fehaciente en grupos homogéneos, según la variable de interés denominada variable de segmentación. Los árboles de decisión constituyen una técnica predictiva ya que la segmentación de la población se realiza según los valores de la variable de interés que juega el papel de variable dependiente del modelo predictivo subyacente en el árbol. La asignación de un elemento poblacional a un segmento se realiza de acuerdo a los valores de determinadas variables medidas sobre él que constituyen las variables independientes del modelo (Molinero, 2002).

Redes neuronales

Por otra parte, las redes neuronales son modelos no lineales, inspirados en el funcionamiento del cerebro, que fueron diseñados para resolver una gran variedad de problemas.

2.17.2. Técnicas descriptivas.

En las técnicas descriptivas no se asignan ningún papel predeterminado a las variables, no se supone la existencia de variables dependientes, ni independientes y tampoco se supone la existencia de un modelo previo para los datos (Pérez López, & Santín Gonzáles, 2007a).

Se pueden mencionar las siguientes técnicas:

Clustering y segmentación

Se trata de buscar agrupamientos naturales en un conjunto de datos tal que tengan semejanzas. Los resultados obtenidos depende de, el algoritmo de agrupamiento seleccionado, el conjunto de datos disponible, la medida de similitud utilizada para comparar objetos (usualmente, definida como medida de distancia) («D3 Clustering.pdf», s. f.).

Jerárquico

El Análisis Clúster Jerárquico (ACJ) es una herramienta exploratoria diseñada para revelar las agrupaciones naturales dentro de un conjunto de datos, los cuales no serían de otra manera evidentes. Los objetos en análisis clúster jerárquico pueden ser casos o variables, dependiendo de si usted desea clasificar casos o examinar relaciones entre las variables («D3 Clustering.pdf», s. f.).

No jerárquico

Inicialmente se establece el número de grupos y cada caso se asigna a uno de ellos.

2.17.3. Algoritmo simple k-means.

Este algoritmo debe definir el número de clusters que se desean obtener, así se convierte en un algoritmo capaz de particionar los datos en grupos correctamente clasificados. Los pasos básicos para aplicar el algoritmo son muy simples, primeramente, se determina la cantidad de clusters en los que se quiere agrupar la información y luego se asume de forma aleatoria los centros por cada clusters.

Una vez encontrados los primeros centroides el algoritmo hará los tres pasos siguientes:

- Determina las coordenadas del centroide.
- Determina la distancia de cada objeto a los centroides.
- Agrupa los objetos basados en la menor distancia.

Finalmente quedarán agrupados por clusters, los grupos de simulaciones según la cantidad de clusters que el investigador definió en el momento de ejecutar el algoritmo («Clustering - EcuRed», 2016).

Técnicas de asociación y dependencia.

Permite predecir patrones de comportamientos futuros sobre ocurrencias simultáneas de valores de variables. Es una técnica no supervisada, en una asociación entre dos atributos ocurre cuando la frecuencia con la que se dan dos o más valores determinados de cada uno conjuntamente es relativamente alta. Las reglas de asociación intentan descubrir asociaciones o conexiones entre objetos (Costas, s. f.).

2.17.4. Técnicas auxiliares.

Son herramientas de apoyo más superficiales y limitadas. Se trata de nuevos métodos basados en técnicas estadísticas descriptivas, consultas e informes y enfocadas en general hacia la verificación.

Proceso analítico de transacciones OLAP

El procesamiento analítico en línea (OLAP, *Online Analytical Processing*) es una tecnología que se usa para organizar grandes bases de datos empresariales y admitir la inteligencia empresarial. Las bases de datos OLAP se dividen en uno o más cubos. Cada cubo lo organiza y diseña un administrador de cubos para que se adapte a la forma en la que recupera y analiza los datos con el fin de que sea más fácil crear y usar los informes de las tablas dinámicas y los gráficos dinámicos que necesite. («Información general sobre el procesamiento analítico en línea (OLAP) - Excel», s. f.).

SQL y herramientas de consulta

SQL Server es la base de la plataforma de datos de Microsoft, ofrece rendimiento crítico con tecnologías en memoria y mejores perspectivas de los datos, ya sea de forma local o en la nube («Microsoft SQL Server», 2016).

2.18. Sistemas de minería de datos

La Figura 6, indica los sistemas de minería de datos más utilizados en el mercado junto con las técnicas que tratan cada uno de ellos, las plataformas sobre las que trabajan y los interfaces de lectura de datos.

Producto	Compañía	Técnicas	Plataforma	Interfaz
Knowledge Seeker	Angoss	Árboles de Decisión	Win	ODBC
CART	Salford Systems	Árboles de Decisión	Win/UNIX	
Clementine	SPSS	Amplio abanico	Win/UNIX	ODBC
Data Surveyor	Data Distilleries	Amplio abanico	UNIX	ODBC
Gain Smarts	Urban Science	Gráficos-Ganancias	Win/UNIX	
Intelligent Miner	IBM	Amplio abanico	UNIX (AIX)	IBM, DB2
Micostrategy	Micostrategy	Datawarehouse	Win	Oracle
Polyanalyst	Megaputer	Simbólicas	Win	Oracle, ODBC
Darwin	Oracle	Amplio abanico	Win/UNIX	Oracle
Enterprise Miner	SAS Institute	Amplio abanico	Win/UNIX/Mac	
SGI MineSet	Silicon Graphics	Asociación y Clasificación	UNIX	Oracle, Sybase, Informix
Wizsoft/Wizwhy	Wizsoft			

Figura 6: Sistemas de minería de datos
Fuente: (Pérez López & Santín Gonzáles, 2007)

2.19. Software estadístico y data mining

IBM SPSS Statistics



Figura 7: Logo SPSS
Fuente: («SPSS», 2016)

Software de análisis estadístico que presenta las funciones principales necesarias para realizar el proceso analítico de principio a fin. Es fácil de utilizar e incluye un amplio rango de procedimientos y técnicas para ayudarle a aumentar los ingresos, superar a la competencia, dirigir investigaciones y tomar mejores decisiones. SPSS Statistics Base proporciona las herramientas básicas de análisis estadístico para cada paso del proceso analítico. Posee una amplia gama de procedimientos estadísticos para dirigir análisis precisos. Contiene técnicas para preparar los datos para el análisis de forma rápida y sencilla. Tiene una funcionalidad completa para elaboración de informes y para la creación de gráficas muy eficientes. Además, sus funciones de visualización sirven para mostrar claramente el significado de sus conclusiones y resultados de las investigaciones. También soporta a todos los tipos de datos, incluidos grandes conjuntos de datos («IBM - SPSS Statistics Base», 2016).

Weka



Figura 8: Logotipo Weka
Fuente: («Mi página de Weka», s. f.)

Escrito en Java, Weka (Entorno Waikato para el Análisis del Conocimiento) es una aplicación de tareas típicas de minería de datos, especialmente pre procesamiento de datos, agrupamiento, clasificación, regresión, visualización y características de selección. Sus técnicas se basan en la hipótesis de que los datos están disponibles en un único archivo plano o relación, donde cada punto marcado es etiquetado por un número fijo de atributos (Ortiz, 2015).

JHepWork



Figura 9: Logotipo JHepWork
Fuente: («image00118.png (844x101)», s. f.)

Diseñado para científicos, ingenieros y estudiantes, jHepWork es un framework para análisis de datos libre y de código abierto que fue creado como un intento de hacer un entorno de análisis de datos usando paquetes de código abierto con una interfaz de usuario comprensible y para crear una herramienta competitiva a los programas comerciales, contiene bibliotecas científicas numéricas implementadas en Java para funciones matemáticas, números aleatorios, y otros algoritmos de minería de datos. jHepWork se basa en Python pero también funciona a la perfección en Java (Ortiz, 2015) .

RapidMiner



Figura 10: Logo RapidMiner
Fuente: («RapidMiner | #1 Open Source Predictive Analytics Platform», s. f.)

Antes llamado YALE, es un ambiente de experimentos en aprendizaje automático y minería de datos que se utiliza para tareas de minería de datos tanto en investigación como en el

mundo real. Permite a los experimentos componerse de un gran número de operadores anidarles arbitrariamente, que se detallan en archivos XML y se hacen con la interfaz gráfica de usuario de RapidMiner. RapidMiner combina esquemas de aprendizaje y evaluadores de atributos del entorno de aprendizaje Weka. Está disponible como una herramienta stand-alone para el análisis de datos y como motor para minería de datos que puede integrarse en tus propios productos («RapidMiner | #1 Open Source Predictive Analytics Platform», 2016).

Orange



Figura 11: Logotipo Orange
Fuente: (Certad, 2010a)

Es un componente de minería de datos y también es un software de aprendizaje de máquina, que permite una programación visual, rápida y versátil para un análisis exploratorio de datos, aunque es una herramienta poderosa, sigue siendo amigable e intuitiva. También permite pre procesamiento, filtros de información, modelación de datos, evaluación de modelos y técnicas de exploración(Certad, 2010b).

CAPITULO III
METODOLOGÍA

3. Generalidades de la investigación

En este capítulo se describen las características y técnicas que se utilizaron para la investigación, considerando los lineamientos metodológicos del proyecto de investigación macro, que busca analizar los usos de la tecnología en los estudiantes de las universidades del Ecuador, basado en los proyectos PIC, DLINHE ECUADOR (Torres, 2012).

La investigación centró como universo de estudio a la Escuela Superior Politécnica de Chimborazo (ESPOCH), ubicada en la provincia de Chimborazo, ciudad de Riobamba, en el Km 1½ de la vía Panamericana Norte; el campus cuenta con las siguientes facultades: Administración de Empresas, Salud Pública, Mecánica, Informática y Electrónica, Ciencias, Recursos Naturales y Ciencias Pecuarias, donde se preparan miles de estudiantes nacionales y extranjeros.



Figura 12: Entrada principal ESPOCH

Fuente: Sitio web http://www.esPOCH.edu.ec/Fotos/noticias/politecnico_de_corazon_ff160.png

3.1. Proceso de extracción del conocimiento a partir de datos (KDD)

Se utilizó como metodología el proceso de extracción del conocimiento a partir de datos KDD, y sus diferentes etapas descritas a continuación.

3.2. Etapa de selección de datos

3.2.1. Fuentes de información

La fuente de información para la investigación fueron todos los estudiantes de la Escuela Superior Politécnica de Chimborazo matriculados en el periodo octubre 2014 – febrero 2015.

3.2.2. Determinación de población y muestra.

Población

La población objeto de estudio fueron los estudiantes del campus Riobamba de la Escuela Superior Politécnica de Chimborazo.

De acuerdo a los datos solicitados en la Dirección de Tecnologías de la Información y Comunicación de la ESPOCH, los datos obtenidos en el reporte del sistema académico OASIS indicaron que en el periodo comprendido desde el 1 de octubre 2014 al 27 de febrero 2015, se encontraban legalmente matriculados 12 865 estudiantes¹.

Muestra

Se realizó un análisis probabilístico para determinar una muestra finita con un nivel de confianza del 95% y un margen de error del 5%. Utilizando la siguiente formula².

$$n = \frac{N \times Z\alpha^2 \times p \times q}{i^2 \times (N - 1) + Z\alpha^2 \times p \times q}$$

Donde:

n: Tamaño de la muestra.

N: Tamaño de la población.

Z α : Coeficiente de confianza, para una distribución de Gauss 0.95% = 1.96

p: Proporción ocurrencia de un evento. (p=0, para mayor tamaño de muestra)

q: Proporción de no ocurrencia de un evento.

i: error

¹ Tomado del Sistema Académico DTIC, ESPOCH. (Noviembre 2014)

² Hernández Sampieri, R., Fernández Collado y Pilar Baptista L. (2006). Metodología de la Investigación. Editorial McGrawHill. México, D.F

Se obtuvo:

- n: ?
- N: 12865
- Z α : 1.96 = 0.05% confianza
- p: 0.5 = 50%
- q: 0.5 = 50%
- i: 0.05 = 5%

Reemplazando:

$$n = \frac{12695 \times 1.96^2 \times 0.5 \times 0.5}{0.05^2 \times (12695 - 1) + 1.96^2 \times 0.5 \times 0.5}$$

$$n = \frac{12192.2780}{31.735 + 0.9604}$$

$$n = 373$$

3.2.3. Obtención de la información

Entrevista

Se realizó una entrevista al Ingeniero Daniel Haro Mendoza, Director del Departamento de Tecnologías de la Información y Comunicaciones de la ESPOCH, quien brindó la ayuda necesaria para recabar la información y respondió a las preguntas planteadas (Ver Anexo 1). Esta entrevista fue grabada y autorizada por el entrevistado para ser publicada en Internet³.

Encuesta

Se aplicó 500 encuestas de forma aleatoria a los estudiantes de la ESPOCH, el modelo de la encuesta utilizada fue proporcionada por el director del proyecto nacional, y consta de 17 preguntas que están enfocadas para determinar las variables sociodemográficas y el uso del Internet en los estudiantes (Ver Anexo 2).

³ Video YouTube, https://youtu.be/1EZf-j_FJ0M

Posteriormente se ingresó cada una de las encuestas en un formulario online mediante la aplicación Survey Monkey⁴, para luego obtener un archivo digital con extensión *.cve para la herramienta de análisis estadístico SPSS⁵.

3.3. Etapa de exploración

Se identificó las variables de la investigación dividiendo en grupos por criterios de semejanza de preguntas, de la siguiente manera:

Distribución de la muestra

Tabla 1: Distribución de la muestra

PREGUNTA	VARIABLE
¿En qué universidad estudia?	V01_SO_UNIVERSIDAD
¿Qué carrera estudia?	V02_SO_CARRERA

Fuente: Elaboración propia

Características sociodemográficas

Tabla 2: Características sociodemográficas

PREGUNTA	VARIABLE
¿Cuál es su edad?	V03_SO_EDAD
¿Cuál es su género?	V04_SO_GENERO
Los ingresos mensuales de su familia son de:	V05_SO_INGRESOS

Fuente: Elaboración propia

Nivel de uso de Internet

Tabla 3: Uso de Internet

PREGUNTA	VARIABLE
¿Desde dónde se conecta habitualmente a Internet? (escoja solo una opción)	V06_INT_CON
De 1 a 7, ¿cuántos días a la semana se conecta Internet?	V07_INT_DIAS
De 1 a 10 su nivel de conocimientos en el manejo de Internet es:	V08_INT_KNOW
¿Aproximadamente cuántas horas se conecta cada día?	V09_INT_HORAS
¿Hace cuántos años se conecta a Internet?	V10_INT_ANIOS

Fuente: Elaboración propia

Uso del Internet para el aprendizaje

Tabla 4: Uso Internet para aprendizaje

PREGUNTA	VARIABLE
¿Cuántas veces a la semana ingresa a la plataforma virtual de su universidad?	V11_EDU_PLAT
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	V12_EDU_PROF
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	V13_EDU_COMP
¿Aproximadamente cuántos recursos educativos descarga de la plataforma virtual cada mes?	V14_EDU_DESC

⁴ Sitio web: <https://es.surveymonkey.com>

⁵ Sitio web: <http://www.ibm.com/analytics/us/en/technology/spss>

¿Aproximadamente cuántos videos académicos mira en YouTube cada mes?	V15_EDU_VIDE
¿Aproximadamente en cuántos foros virtuales participa cada mes?	V16_EDU_FORO
¿Aproximadamente cuántos post o tweets sobre temas académicos realiza en las redes sociales por mes?	V17_EDU_POST
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	V18_EDU_CHAT
¿Aproximadamente cuántas horas busca información académica en internet cada mes?	V19_EDU_INFO
¿Aproximadamente cuántas horas utiliza la biblioteca virtual de la universidad cada mes?	V20_EDU_BIBLI

Fuente: Elaboración propia

Uso del Internet para el entretenimiento

Tabla 5: Uso Internet entretenimiento

PREGUNTA	VARIABLE
¿ Aproximadamente cuántas horas a la semana chatea por diversión?	V21_OCIO_CHAT
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	V22_OCIO_SN
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	V23_OCIO_JUE
¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	V24_OCIO_DESC
¿Aproximadamente cuántos videos para entretenimiento mira en YouTube cada semana?	V25_OCIO_VID

Fuente: Elaboración propia

Interacción en el Internet

Tabla 6: Interacción en Internet

PREGUNTA	VARIABLE
¿ Aproximadamente cuántos seguidores tiene en twitter?	V26_OCIO_TW
¿ Cuántos amigos tiene en Facebook?	V27_OCIO_FB
¿ Cuántos contactos tiene en LinkedIn?	V28_OCIO_LK
Tiene un blog	V29_OCIO_BLOG
Tiene cuenta en YouTube	V30_OCIO_YOU
Tiene cuenta en www.del.icio.us	V31_OCIO_DEL

Fuente: Elaboración propia

Dispositivos usados para conectarse a Internet

Tabla 7: Dispositivos para conectarse a Internet

PREGUNTA	VARIABLE
Smartphone con cámara fotográfica y acceso a Internet	V32_OCIO_SMAR
Teléfono móvil con acceso a Internet	V33_OCIO_CMOVIL
Teléfono móvil sin acceso a Internet	V34_OCIO_SMOVIL
Computador portátil	V35_DEV_PORTATIL
Tablet (iPad; Galaxy Tab, Kindle, etc)	V36_DEV_TABLET
Cámara digital	V37_DEV_CAMDIG
iPod / MP3 Player	V38_DEV_IPOD

Fuente: Elaboración propia

Internet como ayuda en lo académico

Tabla 8: Internet como ayuda en lo académico

PREGUNTA	VARIABLE
Internet le permite elaborar los trabajos más rápido y con menos esfuerzo	V39_INT_MEJOR
Usted confía en la información de Internet para realizar sus tareas	V40_INT_INFO
Internet le permite prescindir de la Biblioteca	V41_INT_BIBLIO
Internet facilita el proceso de aprendizaje	V42_INT_APRE
Internet le permite mejorar sus calificaciones	V43_INT_CALIF
Usted presenta trabajos académicos copiados desde Internet	V44_INT_COPIA

Fuente: Elaboración propia

Usos de herramientas e-learning por los docentes.

Tabla 9: Herramientas e-learning

PREGUNTA	VARIABLE
Su profesor ingresa a la plataforma virtual	V45_PROF_PLAT
Contesta sus consultas por correo electrónico	V46_PROF_MAIL
Chatea con usted eventualmente sobre aspectos académicos	V47_PROF_CHAT
Su profesor comenta en redes sociales sobre temas académicos	V48_PROF_SN
Le sube materiales digitales a la plataforma virtual	V49_PROF_DIGI
Le recomienda recursos digitales de la biblioteca virtual	V50_PROF_RDIGI
Le recomienda videos sobre temas académicos	V51_PROF_VIDEO
Le plantea cuestionarios o evaluaciones en la plataforma virtual	V52_PROF_EVAL
Le plantea foros virtuales	V53_PROF_FORO
Su profesor tiene una página web, blog o perfil de Facebook	V54_PROF_WWW
Su profesor tiene cuenta de twitter	V55_PROF_TW

Fuente: Elaboración propia

Rendimiento académico

Tabla 10: Rendimiento académico

PREGUNTA	VARIABLE
En el semestre anterior, ¿en cuántas asignaturas se matriculó?	V56_SEM_MAT
En el semestre anterior ¿cuántas asignaturas aprobó?	V57_SEM_APRO

Fuente: Elaboración propia

3.4. Etapa de limpieza y transformación

Para la limpieza de datos se realizó las siguientes tareas:

- Estandarizar el nombre de la Universidad.
- Estandarizar los nombres de las carreras, anteponiendo las iniciales de la facultad.
- Corregir errores ortográficos.
- Eliminar las encuestas que tienen como valor 0 en la variable V56_SEM_MAT, puesto que significó que se encuentran en primer semestre.
- Eliminar todas las encuestas que tienen varios datos en blanco.
- Eliminar encuestas incompletas y con datos atípicos.
- Buscar datos en blanco y remplazar por la media.

Al finalizar esta etapa, las 500 encuestas aplicadas se redujeron a 392, que corresponden al número de la muestra (ver página 34), más un 5% adicional que sirve para mejorar el rendimiento de la investigación y disminuir el margen de error.

3.5. Etapa de minería de datos

Se dio tratamiento a los datos, con la búsqueda de patrones utilizando el software estadístico SPSS; se aplicó técnicas de agrupamiento para encontrar características comunes en los grupos de estudio. Este análisis se lo desarrolla ampliamente en el Capítulo IV.

3.6. Etapa de evaluación y presentación de resultados

En el capítulo V, se analiza los resultados obtenidos al aplicar la metodología de minería de datos KDD, presentando los resultados en tablas y graficas estadísticas que sustentan las conclusiones obtenidas.

CAPITULO IV
RESULTADOS

4. RESULTADOS DE LA INVESTIGACIÓN

En el presente capítulo se muestran los resultados obtenidos de las 392 encuestas aplicadas a los estudiantes de la Escuela Superior Politécnica de Chimborazo.

4.1. Distribución de la muestra

De acuerdo a los datos investigados, la población de los estudiantes matriculados en el periodo octubre 2014 a febrero 2015, se distribuyen en las 7 facultades de la siguiente manera: 24% en Administración de Empresas, 15% en Ciencias, 6% en Ciencias Pecuarias, 14% en Informática y Electrónica, 16% en Mecánica, 7% en Recursos Naturales y 19% en Salud Pública.

4.2. Datos Generales

4.2.1. Género

En relación al género, se determina que existe una relación homogénea de género en los encuestados, como se puede observar en la Figura 13.

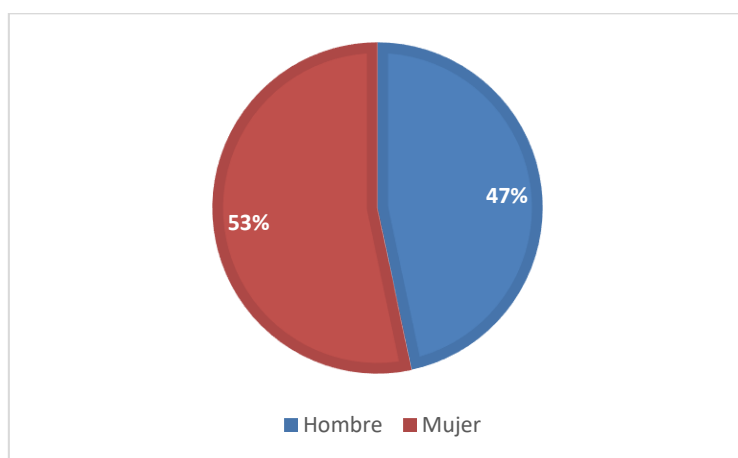


Figura 13: Género
Fuente: Elaboración propia

- Existe una relación directa entre el género con el nivel de ingresos económicos con un nivel de significancia de chi cuadrado ($p=0.008 < 0.05$) donde quienes poseen menores ingresos son las familias de las mujeres.
- Con respecto al lugar de conexión no presenta relación puesto que tiene un nivel de significancia de chi cuadrado ($p=0.644 > 0.05$), es decir no hay diferencia entre el lugar de conexión respecto al género.

- El género tiene una relación directa con el número de horas de conexión a Internet con un valor de significancia de chi cuadrado de ($p=0.026 < 0.05$), es decir que las mujeres se conectan mayor cantidad de horas que los hombres.

4.2.2. Edad

En relación a la edad en la Figura 14, se muestra que el grupo consultado estaba formado por estudiantes entre 18 y 32 años, la edad predominante es de 20 años con 20,9%, la edad máxima es de 32 años, la edad mínima es de 18 años, y la edad promedio de los estudiantes es de 21 a 25 años.

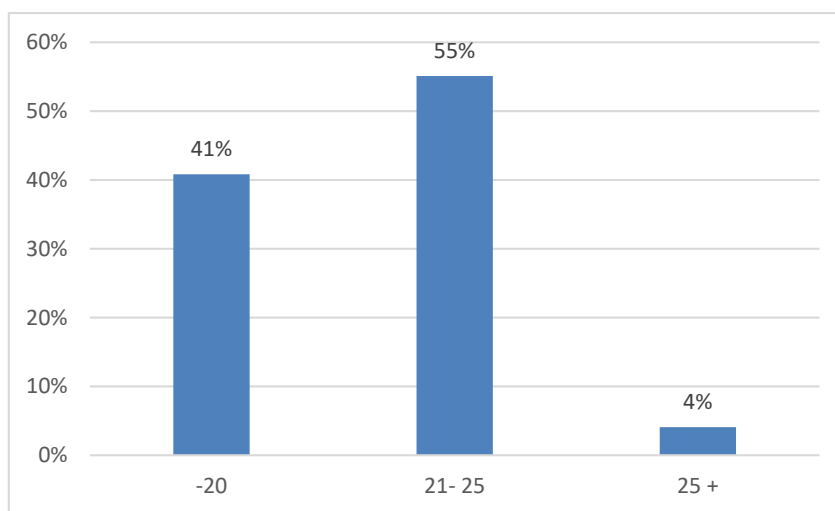


Figura 14: Edad
Fuente: Elaboración propia

- Los resultados muestran que existe una relación directa con el lugar de conexión a Internet ($p=0.000 < 0.05$), donde los encuestados entre 21 a 25 años en mayor proporción se conectan desde la casa.
- En lo referente a las horas de conexión a Internet se tiene una relación directa con significancia de chi cuadrado ($p=0.000 < 0.05$), que indica que los encuestados entre 21 a 25 años se conectan entre 1 a 10 horas diarias.
- La edad, no tiene relación con los videos para entretenimiento con un nivel de significancia ($p=0.982 < 0.05$), pero si tiene una relación directa con los videos académicos con un nivel de significancia ($p=0.000 < 0.05$), es decir que quienes tienen mayor edad miran más videos académicos.

4.2.3. Ingresos

De las familias de los encuestados, el 57% de los encuestados tienen ingresos menores a 600 dólares mensuales, el 34% tiene ingresos entre 600 a 1500 dólares representando un nivel económico medio y solo el 9% de las familias tienen un ingreso mayor a 1500 dólares, lo que claramente indica que el nivel económico de los estudiantes encuestados es bajo (ver Figura 15).

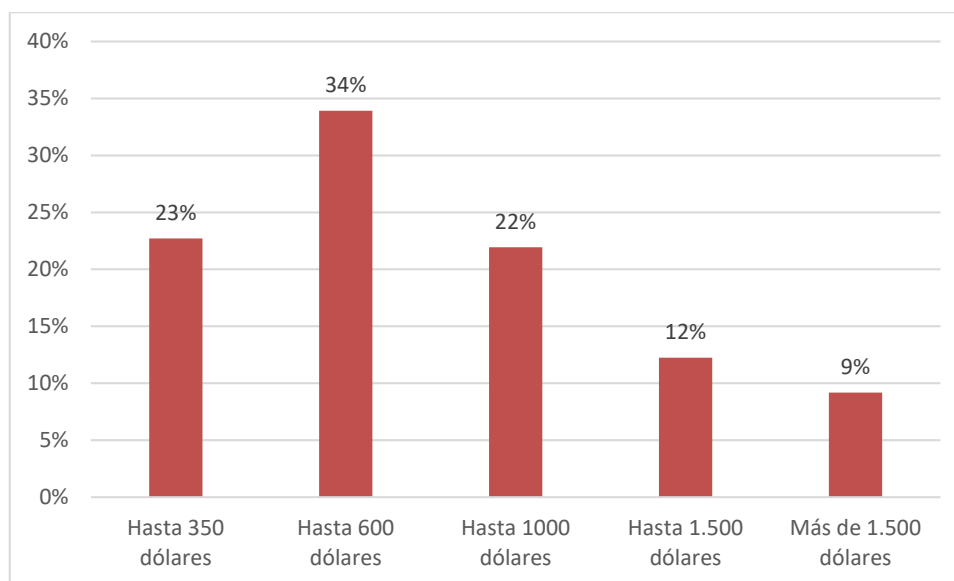


Figura 15: Ingresos mensuales
Fuente: Elaboración propia

- De los datos estudiados se obtiene que existe una relación directa entre el nivel de ingresos con el lugar de conexión a Internet con un nivel de significancia de chi cuadrado ($p=0.000 < 0.05$), es decir que 57% de estudiantes que poseen ingresos menores a 600 se conectan desde la casa.
- Referente a los videos para entretenimiento se tiene un valor significativo (R de Pearson = 0.010), y en relación a los videos académicos se tiene un valor de significancia (R de Pearson = 0.420), lo que indica que hay una relación entre la cantidad de videos que miran para el entretenimiento con los ingresos mensuales, pero no tiene relación con la cantidad de videos que miran los estudiantes para fines académicos.

Con estos datos se verifica que la muestra son estudiantes de la ESPOCH, mayormente entre 19 a 23 años de edad en una proporción casi igual entre hombres y mujeres, en los que el 57% de las familias tienen ingresos menores a 600 dólares.

4.3. Conocimiento y uso del Internet

4.3.1. Preferencias de uso de Internet

Respecto de las preferencias de uso de Internet, el 78% utiliza sus medios de conexión propios siendo el 71% desde la casa y el 7% desde su red móvil, mientras que el 23% lo hacen desde lugares públicos.

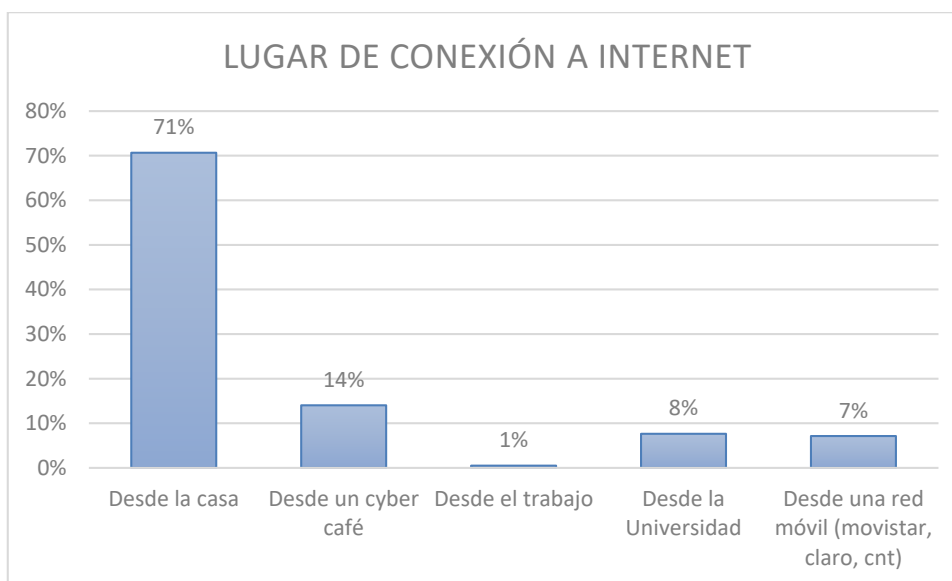


Figura 16: Lugar de conexión a Internet
Fuente: Elaboración propia

En lo referente a los días de conexión a Internet, se determina claramente que el 55% se conectan a Internet todos los días de la semana y el 31% se conectan entre 5 a 6 días, de lo que se concluye que el 86% de los encuestados se conectan más de 5 días por semana al Internet (ver Figura 17).

Los días de la semana en la que los encuestados se conectan a Internet tiene una relación directa con la edad (R de Pearson = 0.006), en la que la mayoría de estudiantes oscilan entre 21 y 25 años, también tiene una débil relación con el género (R de Pearson = 0.047) y una relación directa con los ingresos (R de Pearson = 0.302) de las familias en las que el 57% son de bajos ingresos.

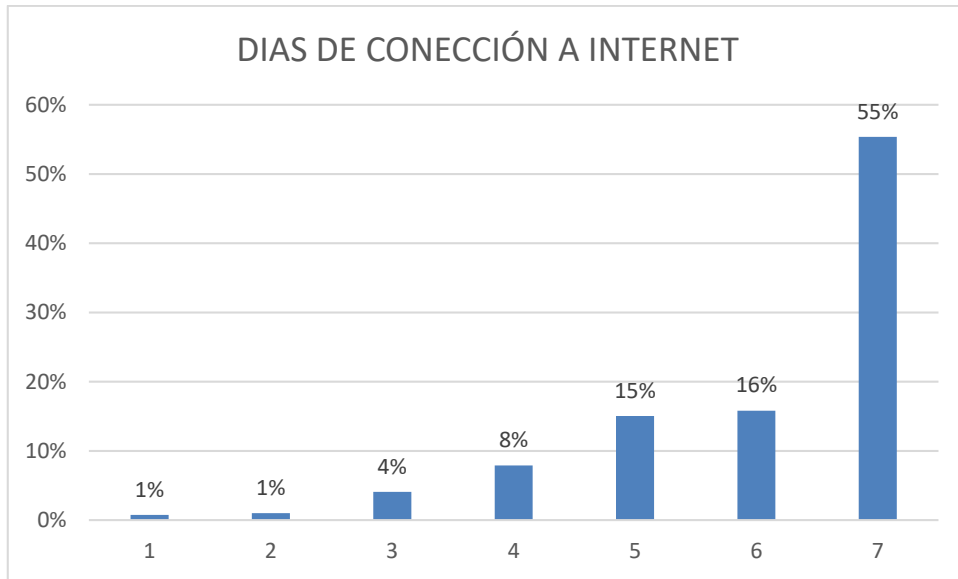


Figura 17: Días de conexión a Internet
Fuente: Elaboración propia

En lo referente al conocimiento en el manejo de Internet el 55% están en la escala entre el 7 al 8 donde los estudiantes poseen un conocimiento medio alto y el 17% se encuentra en un nivel alto, con lo que se determina que el 72% de los encuestados poseen conocimiento en el manejo de Internet, como se muestra en la Figura 18.

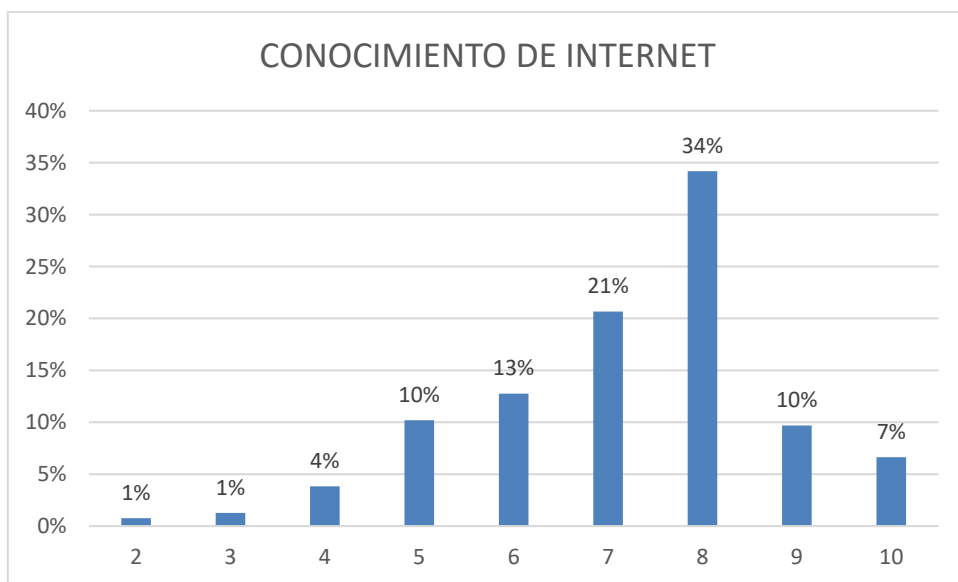


Figura 18: Conocimiento de Internet
Fuente: Elaboración propia

Se determina el tiempo en horas que los encuestados se conectan a Internet por día, encontrando que la mayor concentración de frecuencias va entre 2 a 6 horas diarias, siendo el 68% según la Figura 19.

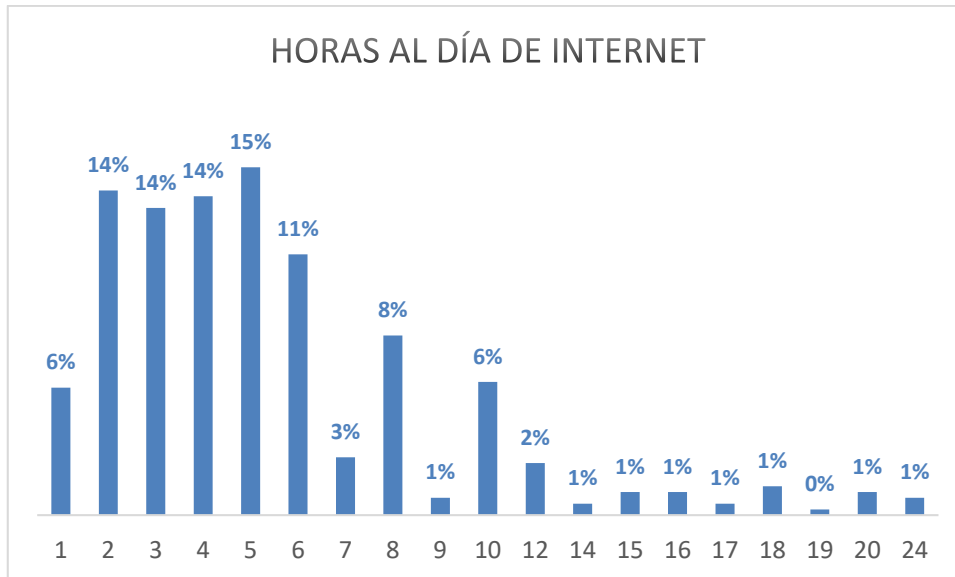


Figura 19: Horas al día de Internet
Fuente: Elaboración propia

De acuerdo a los datos de la Figura 20 se puede apreciar que 91% de los encuestados se conectan a Internet entre 1 a 19 años atrás, considerando que la muestra de estudio tiene entre 19 a 23 años de edad, este dato es muy válido.

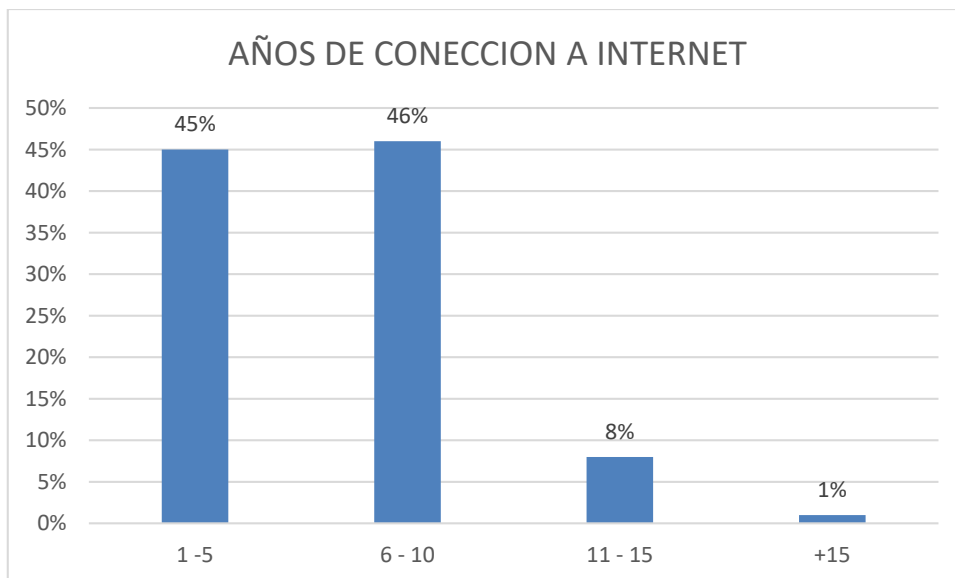


Figura 20: Años de uso de Internet
Fuente: Elaboración propia

Se solicitó a los estudiantes que evaluaran el uso de dispositivos tecnológicos con una valoración de 1 a 10, donde 1 es menos y 10 es más, 48,5% de los encuestados consideraron que el Smartphone con cámara fotográfica y acceso a Internet era útil en una escala de 6 a 10. De igual manera, en una escala de 6 a 10, 53,9% consideró útil el teléfono móvil con

acceso a Internet, 48,2% consideró el celular sin acceso a Internet, un 73,3 expresó la utilidad del computador portátil, 36,7% optó por Tablet. De forma individual por las puntuaciones extremas obtenidas se conoció que 30,6% de los estudiantes no utiliza Smartphone, en comparación con un 17,1% que dice usarlo frecuentemente, en relación al teléfono móvil con acceso Internet de acuerdo a su utilidad 21,9% valoró de menos utilidad frente a un 20,2% que determinó ser de mucha utilidad, sobre el teléfono móvil sin acceso a Internet 27% considero de escasa utilidad contra un 20,4% que expresó resultaba bastante útil, al consultar sobre el computador portátil sólo el 13% no lo consideró lo suficientemente útil, frente al 32,4% que expresó ser una herramienta muy útil, la Tablet no es lo suficientemente útil para el 44,6% de los encuestados y muy útil para un 9,4%, los resultados obtenidos se observan en la Tabla 11.

Tabla 11: Uso de dispositivos

Nivel de Uso	Smartphone	Movil con Internet	Movil sin Internet	Laptop	Tablet
1	30,6%	21,9%	27,0%	13,0%	44,6%
2	3,6%	3,6%	4,3%	1,5%	4,6%
3	4,6%	4,1%	6,4%	1,0%	3,6%
4	5,1%	4,6%	5,4%	3,8%	3,3%
5	7,7%	12,0%	8,7%	7,4%	7,1%
6	6,4%	5,9%	8,4%	6,9%	6,1%
7	5,6%	6,4%	7,7%	8,2%	6,6%
8	12,8%	13,0%	6,1%	13,3%	7,7%
9	6,6%	8,4%	5,6%	12,5%	6,9%
10	17,1%	20,2%	20,4%	32,4%	9,4%

Fuente: Elaboración propia

Con estos datos se puede verificar que 70,7% de los estudiantes habitualmente se conectan desde su casa, el 85% de los estudiantes se conectan más de 5 días a Internet, tienen un promedio entre el 1 a 10 de 7 en conocimiento de manejo del Internet y se conectan en promedio 5 horas al día desde hace 6 años.

4.3.2. Internet para el aprendizaje

Al consultar sobre el número de veces que los encuestados ingresan a la plataforma virtual de la universidad, se observó que 4,4% de los encuestados respondió que no ingresa semanalmente a la plataforma, 16,3% ingresa al menos una vez, 19,6% lo hace dos veces, 19,9% ingresa tres veces, 12,4% ingresa cuatro veces, 11,4% ingresa cinco veces, el 88,8%

de las personas ingresa entre 1 y 9 veces, 5,68% lo hace entre 10 y 19 veces, 0,77% lo hace entre 20 y 29 veces, y solamente 0,25% lo hace un número mayor a cincuenta veces por semana.

Aproximadamente el 88% de los encuestados busca entre 1 y 48 horas al mes información académica, como se muestra en la Figura 21.

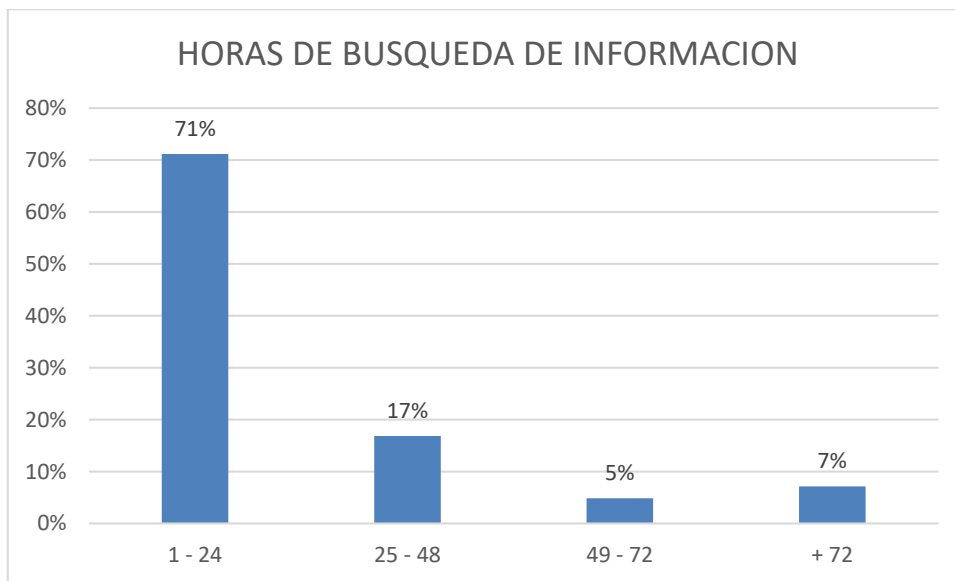


Figura 21 :Horas de búsqueda de información
Fuente: Elaboración propia

Al revisar sobre las características específicas del uso de Internet para fines académicos, se encontró que 1,3% no busca información en Internet, 7,4% no chatea sobre temas académicos, 41,8% no postean o utilizan el tweet para temas académicos, 29,19% no participa en foros virtuales, 20,93% no utiliza las bibliotecas virtuales de la institución, 7,23% no mira foros académicos en YouTube, 14,72% no descarga recursos educativos de la plataforma, 8,01% no realiza consulta ni a sus compañeros ni a sus profesores.

En lo referente al chat académico, los encuestados reportan que el 73% utiliza entre 1 a 10 hora mensuales de Internet, de acuerdo con la Figura 22.

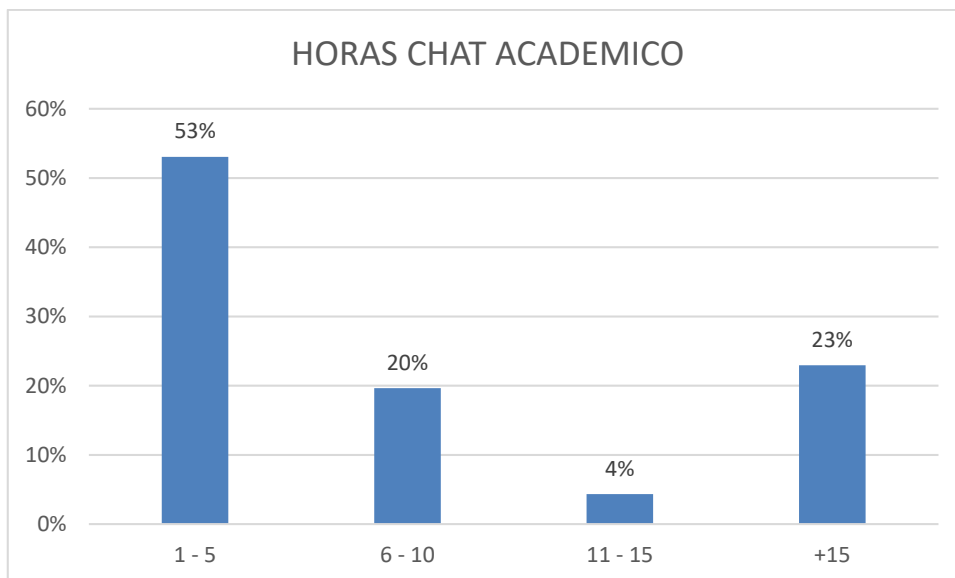


Figura 22 : Horas de chat académico
Fuente: Elaboración propia

Los estudiantes que utilizan el Internet lo hacen entre 1 y más de 65 veces por semana para realizar actividades académicas, es así que 48,83% aseguran realizar consultas a sus maestros entre 1 y 5 veces al mes, 27,13% consulta entre 1 y 5 veces a sus compañeros situaciones académicas, 42,89% descarga entre 1 y 5 recursos por mes, 41,34% mira entre 1 y 5 videos académicos en YouTube, 55,03% participa entre 1 y 5 veces en foros virtuales, 37,46% postea sobre temas académicos entre 1 y 5 veces al mes, 36,69% chatea sobre temas académicos entre 1 y 5 horas al mes, 42,37% utiliza las bibliotecas virtuales de la institución entre 1 y 5 veces, y solo 18,83% realiza búsquedas de información en Internet al mes entre 1 y 5 veces. Adicionalmente se conoció que 0,25% de los consultados declara realizar consultas a sus maestros entre 50 y 55 veces por mes, un 8,01% realiza más de 65 búsquedas de información al mes, 0,5% dice consultarles a sus compañeros más de 65 veces en el mes cuestiones académicas, 2,58% asegura chatear más de 65 horas en el mes sobre temas académicos, 0,75% mira videos académicos más de 65 veces en el mes, 0,25% de los encuestados utiliza las bibliotecas virtuales entre 50 y 55 veces.

De las consultas a sus compañeros de clase el 95% realiza entre 1 a 20 veces al mes de acuerdo a la Figura 23.

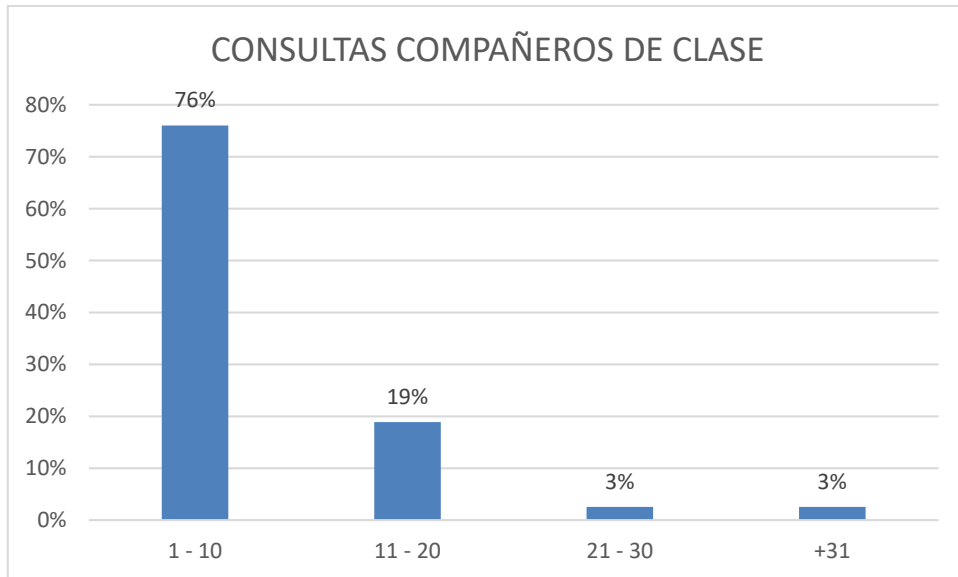


Figura 23 : Consultas académicas a compañeros
Fuente: Elaboración propia

Los estudiantes encuestados, utilizan el Internet principalmente como fuente de Información, luego está el uso del chat académico, seguido por el uso de YouTube en 80% entre 1 a 10 veces al mes (ver Figura 24) para búsqueda de información académica y las consultas a sus compañeros de clases.

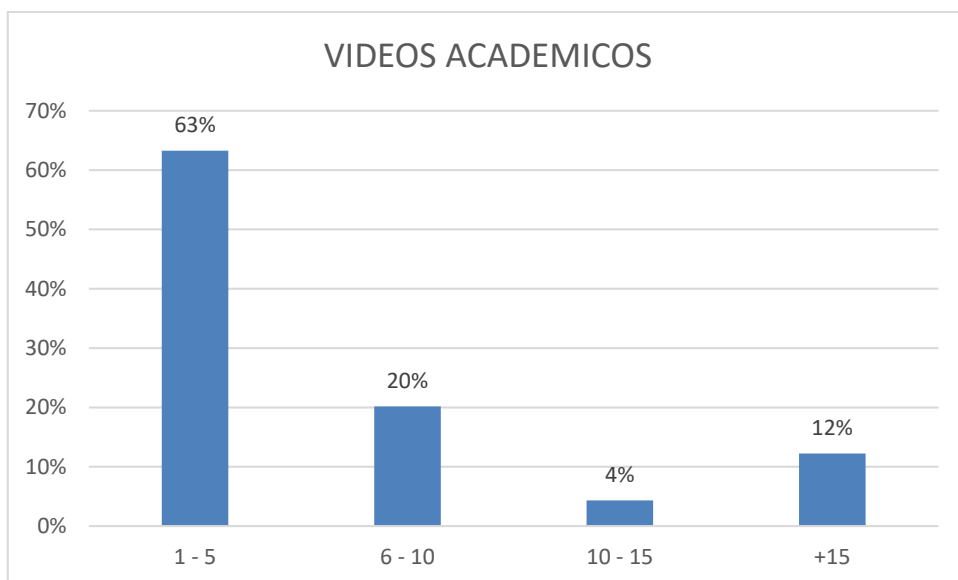


Figura 24 : Videos académicos
Fuente: Elaboración propia

En lo referente a los aspectos de confianza sobre de uso del Internet para el aprendizaje se determinó lo siguiente:

- El 37.0% está en un nivel 8 de acuerdo con que el Internet le permite elaborar los trabajos más rápido y con menos esfuerzo.

- El 20.7% está en un nivel 6 de acuerdo en que confía en la información de Internet para realizar sus tareas.
- El 16.1% está en un nivel 8 de acuerdo con que el Internet le permite prescindir de la biblioteca.
- El 20.9% está en un nivel 10 de acuerdo con que el Internet le facilita el proceso de aprendizaje.
- El 17.1% está en un nivel 8 de acuerdo con que el Internet le permite mejorar sus calificaciones.
- El 15.3% está en un nivel 8 de acuerdo con los encuestados presentan trabajos académicos copiados desde Internet.

Respecto del uso de las herramientas e-learning se determinó que:

- El 70.2% piensa que sus profesores SI, ingresan a la plataforma virtual.
- El 47.4% SI interactúa en consultas con los profesores.
- El 52.6% NO utiliza el chat como medio de comunicación con sus profesores para fines académicos.
- El 49.2% NO utiliza las redes sociales con sus profesores para fines académicos.
- El 78.1% SI utiliza materiales digitales subidos por sus profesores a la plataforma virtual.
- El 68.1% dice que sus profesores SI recomiendan las bibliotecas virtuales.
- El 63.5% dice que sus profesores SI recomienda videos académicos a los estudiantes.
- El 58.9% SI plantean cuestionarios en la plataforma virtual.
- El 56.6% dice que sus profesores SI plantea foros virtuales.
- El 64.8% piensan que sus profesores SI tienen páginas web, blogs o perfil de Facebook.
- El 56.6% piensa que sus profesores NO tienen cuenta de Twitter.

4.3.3. Internet para el entretenimiento

Se identifica que el uso que se le da al Internet en el entretenimiento principalmente se concentra en las redes sociales, luego está el chat, videos en YouTube, descargas de Videos/Música y finalmente los juegos en línea.

En el grupo consultado 40,31% de los encuestados dice no utilizar horas para jugar en línea, 7,23% no lo utiliza para descarga de música, videos y otros, 3% no lo utiliza para mirar videos de entretenimiento en YouTube, 2% no utiliza Internet ni para chatear y el 2% no usa ni para revisar redes sociales.

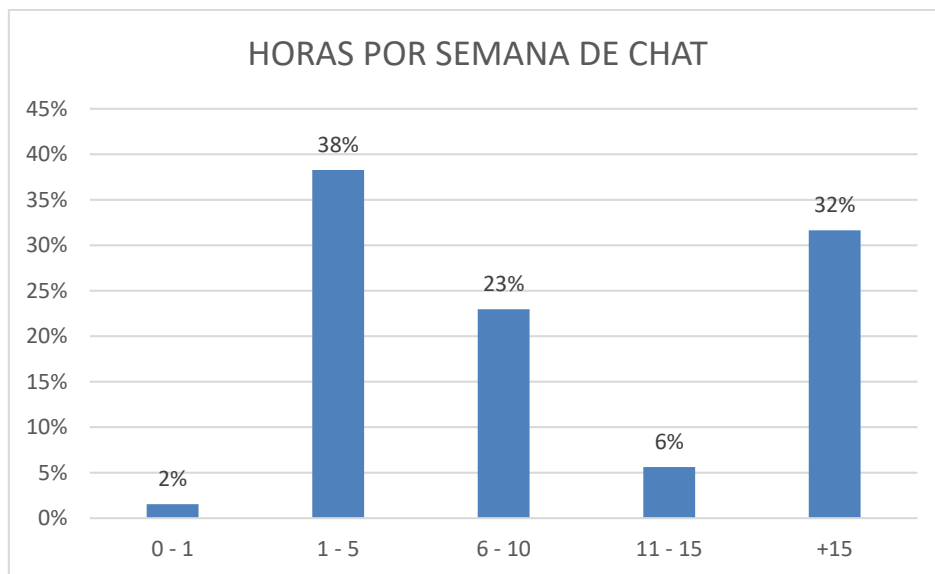


Figura 25: Horas de chat por entretenimiento
Fuente: Elaboración propia

El 49,35% utiliza el Internet entre 1 y 5 veces por semana para descargar música, videos y programas, 37,98% lo utiliza entre 1 y 5 veces por semana para revisar videos de entretenimiento, 29,97% utiliza entre 1 y 5 horas por semana para jugar juegos en línea, 27,90% lo usa entre 1 y 5 horas a la semana para chatear, 21,18% lo utiliza para revisar redes sociales, 3,35% utiliza el Internet más de 65 horas por semana para revisar redes sociales.

En la Figura 26, se muestra que el 33% de los encuestados usa más de 15 horas por semana para el uso de reses sociales.

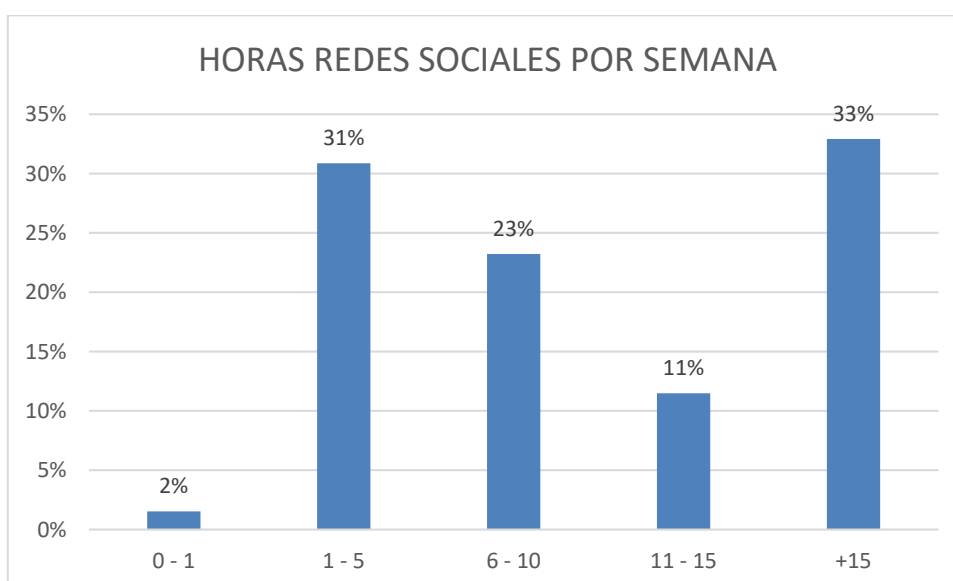


Figura 26: Horas de redes sociales
Fuente: Elaboración propia

Se consultó a los estudiantes que cantidad de amigos y / o seguidores tienen en twitter, y se conoció que 63% no maneja cuenta en Twitter, 32,3% está conectado a entre 1 y 200 personas, 3,1% se conecta con entre 401 y 600 personas. El mínimo de personas es de 1 persona y el máximo es de 1215. De igual manera, se averiguó la situación de amigos o conocidos ligados a un perfil de Facebook, determinándose que 4,4% de personas declara no tener cuenta en Facebook, 27,4% tiene entre 1 y 200 amigos, 25,8% entre 201 y 400 amigos, 15,8% entre 401 y 600 personas, un 9,8% maneja entre 601 y 800 contactos asociados a su perfil, las personas que manejan Facebook, manejan un mínimo de 9 y un máximo de 3000 contactos.

De acuerdo a la Figura 27, se observa que el 22% usa más de 15 horas de videos en línea con fines de entretenimiento y solo el 3% no utiliza Youtube para el entretenimiento.

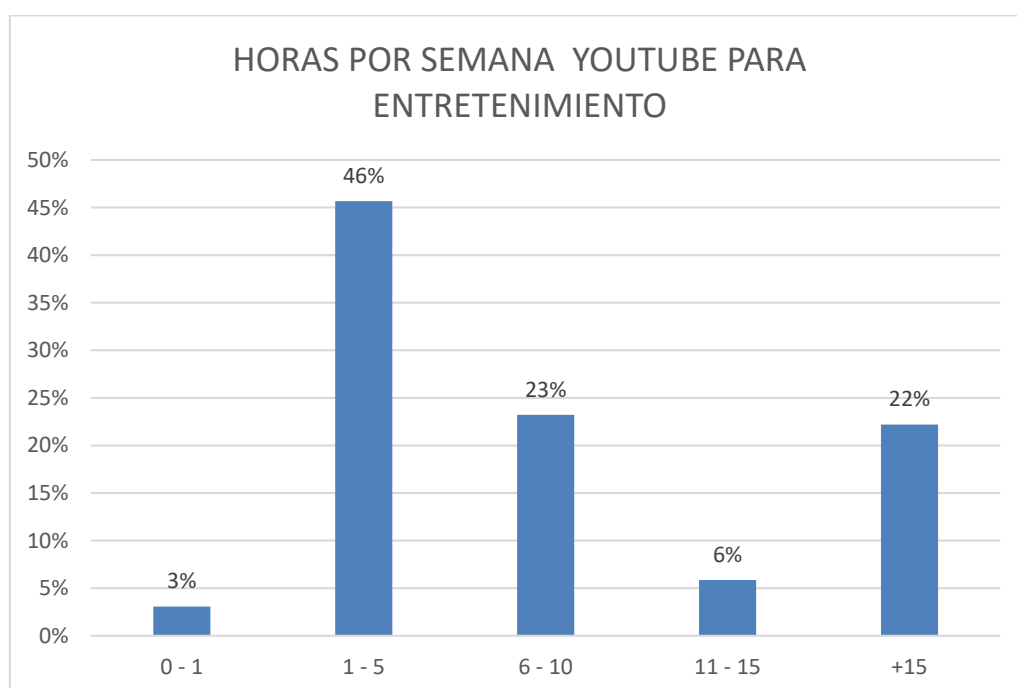


Figura 27 : Horas por semana de videos en Youtube
Fuente: Elaboración propia

4.4. Rendimiento académico

El estudio del rendimiento académico se realiza en base al número de materias que el estudiante se matriculo y el número de materias aprobadas.

Con el fin de mejorar la apreciación de los datos, se crea una variable adicional que representa el número de materias reprobadas (V58_SEM_REPRO) y proviene de la resta del número de materias matriculadas (V56_SEM_MAT) y el número de materias aprobadas (V57_SEM_APRO), con esto se define que, si la diferencia es igual a "0", significa que aprobó todas las materias, y si la variable es diferente de "0", se le asigna "1" para representar estudiante reprobó alguna materia.

Al realizar un análisis estadístico, se puede ver que los estudiantes se matricularon mínimo en 1 materia y máximo en 10, también que mayormente se matriculan en 8 materias, con esto logramos obtener una variable de dos niveles

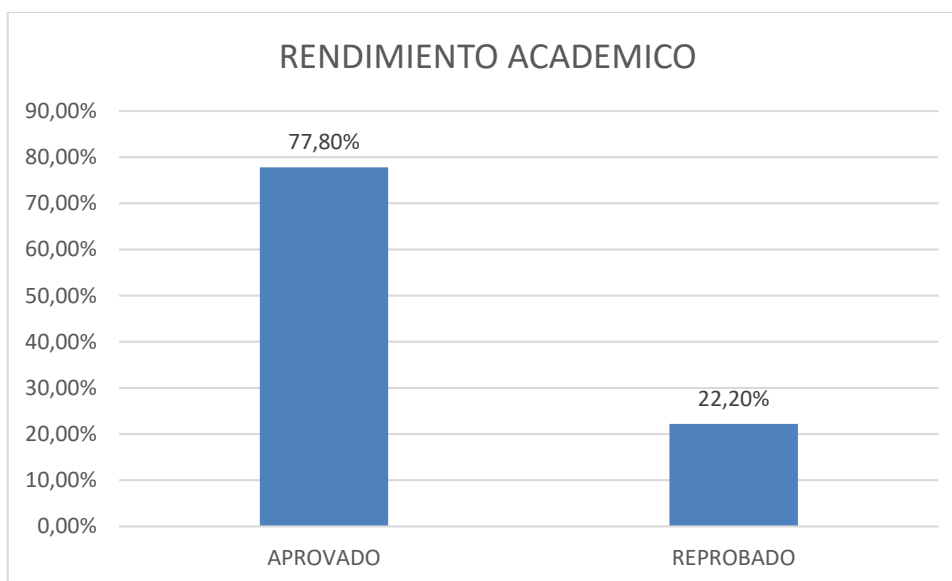


Figura 28 : Rendimiento académico
Fuente: Elaboración propia

Con estos datos se identifica que el 77.8% de los encuestados, no reprobó ninguna materia y el 15% al menos perdió una materia. De los encuestados, el 71.9% se matriculo entre 6 a 8 materias, de los cuales el 67.6% aprobó entre 6 a 8 materias durante el periodo anterior a octubre 2014 – febrero 2015.

4.5. Perfiles de los estudiantes

Para el análisis de los perfiles se realizan los siguientes pasos:

1. Identificación de las variables más representativas de acuerdo al perfil buscado, partiendo del análisis de la encuesta.
2. Clasificar en grupos a los estudiantes, utilizando la técnica de conglomerados k-medias.
3. Realizar el análisis de la discriminante para identificar el número de clusters más adecuado.

4.5.1. Perfil de uso del Internet para el aprendizaje.

Identificación de variables

Las variables más representativas para este análisis son:

- Internet como fuente de Información (Var: V19_EDU_INFO)
- Chat académico (Var: V18_EDU_CHAT)
- Consultas compañeros de clase (Var: V13_EDU_COMP)
- YouTube académico (Var: V15_EDU_VIDE)

Clasificación de grupos

Para clasificar los grupos de estudiantes se utiliza el análisis de conglomerados k-medias, realizando varias pruebas con grupos de 2,3 y 4 clusters (ver Anexo 4). Esta técnica permite agrupar los datos en función de su similitud y formar grupos homogéneos para el estudio de los datos.

Análisis discriminante

Se realiza el análisis discriminante para comprobar e identificar el número de grupos que organiza la información de manera más exacta.

La Tabla 12, indica que el 96.4% de exactitud es para dos grupos que se denominan: *Bajo_Aprendizaje* y *Alto_Aprendizaje*.

Tabla 12 : Exactitud de clasificación

Grupos	Exactitud (%)
De 2	96,40%
De 3	94,10%
De 4	94,10%

Fuente: Elaboración propia

Tabla 13: Número de casos en cada conglomerado

Conglomerado	Alto Aprendizaje	352,000	90%
	Bajo Aprendizaje	40,000	10%
Válidos		392,000	100%
Perdidos		0,000	

Fuente: Software SPSS

El perfil *Alto_Aprendizaje*, está formado por 352 encuestas que corresponde al 90% de la muestra, los cuales utilizan 14 horas de búsqueda de información académica, 8 horas de chat con fines académicos, consultan aproximadamente 8 veces a sus compañeros y miran aproximadamente 6 videos académicos por Internet.

El perfil *Bajo_Aprendizaje*, está formado por 40 encuestas que corresponden al 10% de la muestra, los cuales utiliza 90 horas al mes para búsqueda de información, chatean aproximadamente 46 horas al mes, realizan 19 consultas aproximadamente en el mes y miran aproximadamente 23 videos, todo esto con fines académicos utilizando Internet.

4.5.2. Perfil de uso del Internet para entretenimiento.

Identificación de variables

Las variables más representativas para este análisis son:

- Redes sociales (Var: V22_OCIO_SN)
- Chat entretenimiento (Var: V21_OCIO_CHAT)
- Videos YouTube (Var: V25_OCIO_VID)

Clasificación de grupos

Para clasificar los grupos de estudiantes se utiliza el análisis de conglomerados k-medias, realizando varias pruebas con grupos de 2,3 y 4 clusters (ver Anexo 5). Esta técnica permite agrupar los datos en función de su similitud y formar grupos homogéneos para el estudio de los datos.

Análisis discriminante

Se procede a realizar el análisis discriminante para comprobar e identificar el número de grupos que organiza la información de manera más exacta.

La

Tabla 14, indica que el 95.7% de exactitud es para dos grupos que se los denominan: *Alto_Entretenimiento* y *Bajo_Entretenimiento*.

Tabla 14: Exactitud de clasificación

Grupos	Exactitud (%)
De 2	95,70%
De 3	94,10%
De 4	92,90%

Fuente: Elaboración propia

Tabla 15: Número de casos en cada conglomerado

Conglomerado			
	Alto Entretenimiento	307,000	78%
	Bajo Entretenimiento	85,000	22%
Válidos		392,000	100%
Perdidos		0,000	

Fuente: Software SPSS

El perfil *Alto_Entretenimiento*, está formado por 307 encuestas correspondientes al 78% de la muestra, de los cuales aproximadamente usan 9.2 horas semanales las redes sociales, 38.6 horas semanales de chat por diversión y miran 7.7 videos para entretenimiento cada semana.

El perfil *Bajo_Entretenimiento*, está formado por 85 encuestas que corresponden al 22% de la muestra, quienes aproximadamente utilizan 41 horas semanales en redes sociales, 38

horas semanales en chat por diversión y miran 22 videos para entretenimiento por semana.

4.6. Pruebas de hipótesis

Para la comprobación de hipótesis se realiza un análisis de regresión logística binaria para la comprobación del modelo estadístico, debido a que la variable dependiente tiene dos niveles.

4.6.1. Hipótesis 1.

Hipótesis 1: El nivel de ingresos económicos determina como se utiliza Internet para el aprendizaje.

Variable independiente: El nivel de ingresos.

Variable dependiente: Usos del Internet para el aprendizaje.

Ho: El nivel de ingresos no incide en el uso del Internet para el aprendizaje.

H1: El nivel de ingresos si incide en el uso del Internet para el aprendizaje.

Regresión logística binaria

En el análisis de regresión logística binaria, se considera como variable binaria al perfil de uso del Internet para el aprendizaje considerando: *Bajo_Aprendizaje* el (10% de la muestra) y *Alto_Aprendizaje* (90% de la muestra).

En la Tabla 16: Tabla de clasificaciónTabla 16, se verifica que hay un 89.8% de casos correctamente clasificados concentrados en el conglomerado de Alto Aprendizaje. Con estos datos se acepta el modelo puesto que más del 50% de los datos están correctamente clasificado.

Tabla 16: Tabla de clasificación

Observado		Pronosticado			
		Número inicial de casos		Porcentaje correcto	
		Alto Aprendizaje	Bajo Aprendizaje		
Paso 0 y Paso 1	Número inicial de casos	Alto Aprendizaje	352	0	100,0
		Bajo Aprendizaje	40	0	0,0
Porcentaje global					89,8

a. En el modelo se incluye una constante.

b. El valor de corte es ,500

Fuente: Elaboración propia

En la Tabla 17, se demuestra que el valor de significancia es mayor que 0.05 de lo que se deduce que la variable que se está introduciendo no aporta a mejorar la probabilidad de incidencia de la variable dependiente.

Tabla 17: Variables que no están en la ecuación

			Puntuación	gl	Sig.
Paso 0	Variables	Hasta 350 dólares	3,789	4	,435
		Hasta 600 dólares	1,507	1	,220
		Hasta 1000 dólares	1,460	1	,227
		Hasta 1.500 dólares	,804	1	,370
		Más de 1.500 dólares	,209	1	,648
	Estadísticos globales		3,789	4	,435

Fuente: Elaboración propia

En la prueba Ómnibus en la Tabla 18, permite medir la significancia del modelo, se verifica que el valor de significancia (0.408) es mayor que 0.05, es decir que el modelo no es significativo, por lo tanto, las variables independientes no explican a la variable dependiente. De la misma forma el valor de R^2 de Nagelkerke es 0.021, es muy bajo para aceptar el modelo (Anexo 7).

Tabla 18: Pruebas ómnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	3,988	4	,408
	Bloque	3,988	4	,408
	Modelo	3,988	4	,408

Fuente: Elaboración propia

Con la prueba de Hosmer y Lemeshow en la Tabla 19, se indica que el valor de chi-cuadrado no es significativo ($X^2 = 0.000$, $p = 1.000$), con lo que se puede indicar que el modelo si se ajusta a los datos con esta prueba.

Tabla 19: Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	3	1,000

Fuente: Elaboración propia

El valor de significancia de las variables de la ecuación en la Tabla 20, tienen un valor de significancia superior a 0.05, con lo que demuestra que la variable independiente no aporta para mejorar la predicción de la variable dependiente.

Tabla 20: Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Paso 1^a	Hasta 350 dólares		3,656	4	,455			
	Hasta 600 dólares	,206	,842	,060	1	,806	1,229	,236 6,395
	Hasta 1000 dólares	,913	,773	1,396	1	,237	2,491	,548 11,325
	Hasta 1.500 dólares	,914	,796	1,317	1	,251	2,493	,524 11,867
	Más de 1.500 dólares	,435	,896	,236	1	,627	1,545	,267 8,942
	Constante	-2,833	,728	15,162	1	,000	,059	
	a. Variable(s) introducida(s) en el paso 1: V005_INGRESOS.							

Fuente: Elaboración propia

Considerando que el modelo de regresión logística binaria es válido, puesto que se clasifica correctamente el 89.8% de los casos, pero que en la prueba de bondad no es significativo respecto a la predicción de la variable dependiente y que las variables de la ecuación no son significativas ($p < 0.05$), por lo tanto, no hay incidencia del nivel de ingresos económicos sobre el uso del Internet para el aprendizaje, de tal manera que, **se acepta la hipótesis nula y se rechaza la hipótesis alternativa.**

4.6.2. Hipótesis 2.

Hipótesis 2: El nivel de ingresos económicos determina como se utiliza Internet para el entretenimiento.

Variable Independiente: El nivel de ingresos.

Variable dependiente: Usos del Internet para el entretenimiento.

Ho: El nivel de ingresos no incide en el uso del Internet para el entretenimiento.

H1: El nivel de ingresos si incide en el uso del Internet para el entretenimiento.

Regresión logística binaria

En este análisis se considera los dos conglomerados del perfil de uso del Internet para el entretenimiento: *Alto_Entretenimiento* (78% de la muestra) y *Bajo_Entretenimiento* (22% de la muestra).

En los pasos 0 y paso 1, se puede ver en la Tabla 21, que existe el 78.3% de los casos correctamente clasificados, con estos datos se acepta el modelo puesto que más del 50% de todos los casos están correctamente clasificados.

Tabla 21: Tabla de clasificación

Observado			Pronosticado		
			Número inicial de casos		Porcentaje correcto
			1	2	
Paso 0 y Paso 1	Número inicial de casos	1	307	0	100,0
		2	85	0	0,0
Porcentaje global					78,3
a. En el modelo se incluye una constante.					
b. El valor de corte es ,500					

Fuente: Elaboración propia

En la Tabla 22, se verifica que el valor de significancia para las dos primeras variables (0.017 y 0.015) son menores que 0.05 lo que implica que si aportan en la predicción de la variable dependiente.

Tabla 22: Variables que no están en la ecuación

			Puntuación	gl	Sig.
Paso 0	Variables	Hasta 350 dólares	12,014	4	,017
		Hasta 600 dólares	5,895	1	,015
		Hasta 1000 dólares	1,569	1	,210
		Hasta 1.500 dólares	2,513	1	,113
		Más de 1.500 dólares	1,804	1	,179
	Estadísticos globales		12,014	4	,017

Fuente: Elaboración propia

En la prueba Ómnibus de la Tabla 23, se encuentra que el valor de significancia (0.016) es menor que 0.05 por lo tanto esto quiere decir que la introducción de la variable si mejora en la predicción de la variable dependiente (Anexo 8).

Tabla 23: Pruebas Ómnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	12,170	4	,016
	Bloque	12,170	4	,016
	Modelo	12,170	4	,016

Fuente: Elaboración propia

Mediante la prueba de Hosmer y Lemeshow en la Tabla 24, se indica que el valor de chi-cuadrado no es significativo ($X^2 = 0.000$, $p = 1.000$), con lo que se puede decir que el modelo si se ajusta a los datos con esta prueba y ayuda a mejorar la incidencia a la variable dependiente.

Tabla 24: Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	3	1,000

Fuente: Elaboración propia

En la Tabla 25, el valor de significancia de las variables de la ecuación, en las dos primeras variables tiene un valor menor a 0.05, lo que implica que si ayuda a mejorar los valores de predicción de la variable dependiente.

Tabla 25: Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1^a	Hasta 350 dólares			11,535	4	,021			
	Hasta 600 dólares	-1,266	,478	7,004	1	,008	,282	,110	,720
	Hasta 1000 dólares	-,820	,419	3,825	1	,050	,440	,194	1,002
	Hasta 1.500 dólares	-,256	,428	,358	1	,549	,774	,335	1,790
	Más de 1.500 dólares	-,194	,475	,167	1	,683	,824	,324	2,090
	Constante	-,693	,354	3,844	1	,050	,500		
a. Variable(s) introducida(s) en el paso 1: V005_INGRESOS.									

Fuente: Elaboración propia

Analizando los ODDs, se encuentra que los Exp(b) son menores de 1 para los casos significativos, lo que señala que un aumento de la variable independiente, disminuye los ODDs que ocurra el evento, por lo tanto, es probable que el aumento del nivel de ingresos de las familias con menos de 600 dólares mensuales, disminuya la probabilidad de la incidencia del uso del Internet para el entretenimiento, y si disminuye los ingreso es menos probable que suceda el evento (Signo – en B).

Considerando que el modelo de regresión logística binaria es válido, y que el nivel de significancia para los casos en que los ingresos son menores que 600 dólares es significativo ($p < 0.05$), se determina que, si existe incidencia para los casos menores a 600 dólares, pero no existe incidencia entre los ingresos mayores, por lo tanto, **se acepta la hipótesis nula y se rechaza la hipótesis alternativa**, debido a que en el análisis de los ODD indica que a mayores ingresos hay menor probabilidad de incidencia de las variables.

4.6.3. Hipótesis 3.

Hipótesis 3: El uso de la tecnología en el aprendizaje incide en el rendimiento académico.

Variable Independiente: Usos de la tecnología en el aprendizaje

Variable dependiente: Rendimiento académico

Ho: El uso de la tecnología en el aprendizaje no incide en el rendimiento académico

H1: El uso de la tecnología en el aprendizaje si incide en el rendimiento académico

Para este análisis, se utiliza el perfil de usos de Internet para el aprendizaje considerando que en la encuesta la tecnología que se estudia es el acceso a Internet para fines educativos y de entretenimiento, y que el 32% de la muestra utiliza como dispositivo tecnológico la laptop para conectarse a la red.

Regresión logística binaria

Utilizando la variable de materias aprobadas V58_SEM_REPRO, se considera como una variable dicotómica debido a que tiene dos estados "0" cuando el estudiante aprobó todas las materias y "1" cuando el estudiante reprobó una o más materias.

Mediante la Tabla 26: Tabla de clasificaciónTabla 26 de clasificación, se puede validar que existe un 77.8% de datos correctamente clasificados, por lo tanto, se acepta el modelo debido a que más del 50% de los casos están correctamente clasificados.

Tabla 26: Tabla de clasificación

Observado			Pronosticado		
			Materias Reprobadas V56-V57		Porcentaje correcto
			,00	1,00	
Paso 0	Materias Reprobadas V56-V57	,00	305	0	100,0
		1,00	87	0	0,0
	Porcentaje global				
a. En el modelo se incluye una constante.					
b. El valor de corte es ,500					

Fuente: Elaboración propia

En la Tabla 27, se muestra que el valor de significancia para las variables que no están en la ecuación es mayor que 0.05 con lo que se deduce que la variable no existe incidencia en la variable dependiente.

Tabla 27: Variables que no están en la ecuación

		Puntuación	gl	Sig.	
Paso 0	Variables	QCL_AP_2(1)	,726	1	,394
	Estadísticos globales		,726	1	,394

Fuente: Elaboración propia

La Tabla 28, muestra los resultados de la prueba Ómnibus en la que se puede ver que el valor de significancia (0.405) es mayor que 0.05 por lo que se entiende que la introducción de la variable independiente no incide en la variable dependiente y el modelo no es significativo. También el valor de R² de Nagelkerke es 0.003, que es muy baja para aceptar el modelo (Anexo 9).

Tabla 28: Pruebas Ómnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	,694	1	,405
	Bloque	,694	1	,405
	Modelo	,694	1	,405

Fuente: Elaboración propia

El valor de significancia de las variables de la ecuación en la de la Tabla 29, indican que son mayores a 0.05 por lo que no hay incidencia en la mejora de la predicción de la variable dependiente.

Tabla 29: Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)		
							Inferior	Superior	
Paso 1^a	QCL_AP_2(1)	-,320	,377	,721	1	,396	,726	,347	1,520
	Constante	-,969	,354	7,494	1	,006	,379		
a. Variable(s) introducida(s) en el paso 1: QCL_AP_2.									

Fuente: Elaboración propia

Considerando que el modelo de regresión logística binaria es válido, puesto que se clasifica correctamente el 77.8% de los casos, pero que en la prueba de bondad no es significativo respecto a la predicción de la variable dependiente y que las variables de la ecuación no son significativas ($p > 0.05$), entonces, no hay incidencia del uso de la tecnología en el aprendizaje sobre el rendimiento académico de los estudiantes. Por lo tanto, **se acepta la hipótesis nula y se rechaza la hipótesis alternativa.**

4.6.4. Hipótesis 4.

Hipótesis 4: El uso de la tecnología para entretenimiento incide en el rendimiento académico.

Variable Independiente: Usos de la tecnología para el entretenimiento.

Variable dependiente: Rendimiento académico.

Ho: El uso de la tecnología para entretenimiento no incide en el rendimiento académico.

H1: El uso de la tecnología para entretenimiento si incide en el rendimiento académico.

Para este análisis, se utiliza el perfil de usos de Internet para el entretenimiento considerando que en la encuesta la tecnología que se estudia es el acceso a Internet para fines educativos y de entretenimiento, y que el 32% de la muestra utiliza como dispositivo tecnológico la laptop para conectarse a la red.

Regresión Logística Binaria

Utilizando la variable de materias aprobadas V58_SEM_REPRO, se considera como una variable dicotómica debido a que tiene dos estados "0" cuando el estudiante aprobó todas las materias y "1" cuando el estudiante reprobó una o más materias.

Tabla 30: Tabla de clasificación

Observado			Pronosticado		
			Materias Reprobadas V56-V57		Porcentaje correcto
			,00	1,00	
Paso 0	Materias Reprobadas V56-V57	,00	305	0	100,0
		1,00	87	0	0,0
	Porcentaje global				77,8
a. En el modelo se incluye una constante.					
b. El valor de corte es ,500					

Fuente: Elaboración propia

Con los datos de la Tabla 31, se muestra que los valores de significancia de las variables (0.254) es mucho mayor que 0.05 por lo que se determinan que no existe incidencia en la variable dependiente.

Tabla 31: Variables que no están en la ecuación

		Puntuación	gl	Sig.
Paso 0	Variables	1,299	1	,254
	Estadísticos globales	1,299	1	,254

Fuente: Elaboración propia

En la prueba Ómnibus de la Tabla 32, se observa que el nivel de significancia es mayor a 0.05 por lo tanto entendemos que al introducir la variable independiente no se mejora la predicción de la variable dependiente. También el valor de R^2 de Nagelkerke es 0.005, que es muy baja para aceptar el modelo (Anexo 10).

Tabla 32: Pruebas omnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	1,352	1	,245
	Bloque	1,352	1	,245
	Modelo	1,352	1	,245

Fuente: Elaboración propia

Al revisar la Tabla 33, se verifica que el valor de significancia para las variables (0.256) es mayor que 0.05 por lo tanto verificamos que no ayuda a mejorar la predicción de la variable dependiente.

Tabla 33: Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1^a	QCL_EN_2	-,358	,315	1,290	1	,256	,699	,377	1,296
	Constante	-,825	,392	4,437	1	,035	,438		
a. Variable(s) introducida(s) en el paso 1: QCL_EN_2.									

Fuente: Elaboración propia

Analizando los datos obtenidos, se determina que el modelo de regresión logística binaria es válido, puesto que se clasifica correctamente el 77.8% de los casos, pero que en la prueba de bondad no es significativo respecto a la predicción de la variable dependiente y que las variables de la ecuación no son significativas ($p > 0.05$), entonces, no hay incidencia del uso de la tecnología en el aprendizaje sobre el rendimiento académico de los estudiantes. Por lo tanto, **se acepta la hipótesis nula y se rechaza la hipótesis alternativa**

CAPITULO V
DISCUSIÓN DE RESULTADOS

5. Discusión de resultados

En este capítulo se discute los resultados obtenidos en la investigación, respecto a la comprobación de las hipótesis planteadas, para cumplir con los objetivos propuestos y proponer trabajos futuros.

5.1. Perfil de los estudiantes

En la exploración de los datos, se verifica la muestra está distribuida entre las 7 facultades, con 24% en Salud Pública, 21% en Mecánica, 18% en Informática y Electrónica, 14% en Ciencias, 13% en Administración de Empresas, 5% en Recursos Naturales y 5% en Ciencias Pecuarias, estos datos tienen una distribución proporcional a la población.

El rango de edades de los encuestados oscila en un 76.8% entre los 19 y 23, de los cuales el 53% son mujeres y el 47% son hombres, y en lo referente a los ingresos económicos el 56.6% de sus familias tienen ingresos menores a 600 dólares por mes, por lo que podemos decir que el perfil socioeconómico de los encuestados se define como: hombres y mujeres de entre 19 a 23 años de edad con un nivel de ingresos familiar bajo menor a 600 dólares por mes.

A un cuando el nivel económico es bajo, el 55.4% de encuestados se conectan a Internet todos los días, principalmente desde la casa, en un promedio de 5.6 horas por día, aproximadamente desde hace 6 años atrás, y en una escala del 1 al 10 consideran que tienen un conocimiento medio-alto nivel 7 de manejo del Internet.

Se obtuvo que el 55% de los encuestados se conecta los siete días en la semana, 16% lo hace seis días, 15% se conecta cinco días, 8% lo hace cuatro días, 4,1% lo hace tres días, 1% lo hace dos días, y solamente 0,8% utiliza el Internet un día a la semana, en relación al género, en base a las desviaciones típicas entre hombres (1,26) y mujeres (1,37), se verifica que los grupos son homogéneos en las conexiones a Internet durante la semana, los resultados discrepan de lo reportado en el estudio sobre el uso de las TIC en estudiantes de Psicología de la Universidad Católica del Maule, en donde al medir el rendimiento académico, género, edad, antigüedad, con el uso de las TIC durante el proceso de enseñanza y aprendizaje, valoración de la plataforma virtual, valoración del correo electrónico, uso de la plataforma virtual, buscadores y revistas electrónicas, valoración del uso de multimedia, entre otros, determinaron que a pesar de que la gran mayoría de estudiantes utiliza las TIC con

fines académicos, hay diferencias marcadas de género con un 57% de los varones usa la computadora todos los días, en comparación con un 37% de las mujeres (Muñoz, 2008), mientras que estos resultados concuerdan con Torres, donde indica que el género no tiene asociación con los perfiles académicos sobre el uso de Internet (Torres, 2012).

En relación al nivel de manejo de Internet, los estudiantes que contestaron la encuesta se autocalificaron entre buenos y muy buenos, es así que 54,5% de los encuestados consideran que se encuentran entre 7 y 8 puntos al valorarse y 16,3% se autocalifica entre 9 y 10, las calificaciones más altas, los resultados obtenidos son similares a los de un estudio similar sobre la búsqueda de información con fines académicos entre el alumnado universitario, en donde se determinó que 46,2% se consideraba bueno en el manejo de internet y 23,2% se autocalificaba muy bueno (Comas, Sureda, Pastor, & Morey, 2011).

Con respecto al acceso, 70,5% de los encuestados aseguró tener Internet en sus hogares y conectarse desde allí, 14% se conecta desde un cyber café, 8% lo hace desde la Universidad, 7% lo hace desde una red móvil de cualquier operadora y solamente un 0,5% de los encuestados asegura hacerlo desde el lugar en donde trabaja, estos datos difieren notablemente de los presentados por Muñoz en 2008, donde se asegura que en el manejo de TIC de los estudiantes de la Universidad del Maule en Chile, un 46% utiliza el computador de su casa y un 47% accede al computador en la universidad, siendo las mujeres quienes más utilizan el computador de la universidad (50%) en comparación a los varones (42%). Las causas de las diferencias pueden encontrarse en la amplia difusión de Internet domiciliario, que incide para que un número mayor de familias accedan a Internet desde sus hogares y al tiempo que ha transcurrido desde el 2008 hasta la actualidad en la que se ha masificado el acceso a Internet desde el hogar.

En la relación entre el género y el nivel de ingresos, mediante el uso del software SPSS y aplicando tablas de contingencia, el valor de significancia de chi-cuadrado indica que existe relación entre las mujeres y los menores ingresos económicos; respecto al lugar de conexión no presenta diferencia, pero al relacionar con el número de horas de uso de Internet, si se encuentra que las mujeres se conectan mayor cantidad de horas a Internet con relación a los hombres.

Los resultados muestran que la edad de los estudiantes, tiene una relación directa con el lugar de conexión en la cual, los encuestados que están en un rango de edad entre los 21 a 25 años indican que mayormente se conectan desde su casa, entre 1 a 10 horas al día. Esto se

comprueba mediante la aplicación de una tabla de contingencia en donde el valor de significancia de chi-cuadrado presenta una relación significativa.

Mediante tablas de contingencia, al estudiar la edad y su relación con el tipo de videos que miran los estudiantes (entretenimiento o académicos), se verifica que no existe relación significativa ($p > 0.05$) con los videos para entretenimiento, pero si existe relación significativa ($p < 0.05$) con los videos para el ámbito académico, lo que indica que existe dependencia entre la edad y la cantidad de videos académicos que miran los estudiantes por Internet, lo que significa que la edad es un factor importante en el uso del Internet, considerando que los estudiantes de mayor edad invierten mejor el tiempo en el uso de Internet para fines académicos.

Los ingresos económicos determinan las posibilidades de acceso a la tecnología, es decir a tener una conexión de Internet y tener uno o varios dispositivos para acceder a la red, lo que es un factor de incidencia directa con el lugar de conexión donde el 57% de los estudiantes indican que tienen ingresos menores a 600 dólares mensuales, y que se conectan desde la casa, estos resultados coinciden con la investigación de (Zaragozión, 2016).

5.2. Usos del Internet

Para el **perfil del uso del Internet para el aprendizaje**, las variables más representativas consideradas para el modelo son: el Internet como fuente de información, el chat académico, las consultas a los compañeros de clase y YouTube académico, con las cuales se determinó dos conglomerados, *Alto_Aprendizaje* representado por el 90% de la muestra y *Bajo_Aprendizaje* con el 10% de la muestra.

Los principales usos que le dan al Internet en lo académico es la búsqueda de información, chat, consultas a sus compañeros y videos académicos en YouTube; se identifica que el 40% de los encuestados tienen un blog, y el 70% de ellos tienen una cuenta en YouTube. El 37% está de acuerdo en que Internet le permite elaborar los trabajos más rápido y con menos esfuerzo, el 20.7% confía en toda la información que encuentra y el 15% acepta haber copiado trabajos que ha presentado en clases desde Internet.

En relación al uso de TICs durante el proceso de enseñanza se encontró que cada mes, el 91,98% realiza consultas a sus profesores o a sus compañeros, 85,27% descarga recursos educativos de la plataforma virtual, 92,76% mira videos en YouTube, 70,80% participa en

foros virtuales, 58,13% postea temas académicos, 92,50% chatea sobre temas relacionados a sus estudios, 98,70% busca información en Internet con fines educativos y 79,06% utiliza la biblioteca virtual, estos datos son superiores a los reportados por Muñoz en 2008, que asegura que los estudiantes en 30% utilizó la plataforma virtual, un 24% utilizó el correo electrónico, un 5,5% utilizó buscadores web y revistas electrónicas, un 35% menciona que todas las anteriores y un 4% menciona que ninguna, las diferencias pueden radicar en el rápido crecimiento y el auge del Internet en la última década.

En lo referente al uso de las herramientas e-learning, los encuestados indican que el 70% de los profesores utilizan el aula virtual, el 78% utiliza material publicado en la plataforma virtual y el 68% utiliza las bibliotecas virtuales, los estudiantes aseguran que los docentes utilizan tecnologías de la información y la comunicación, 70% de los docentes ingresa a la plataforma virtual, 78% de los docentes sube materiales digitales al aula virtual, 67,7% les recomienda recursos digitales de la biblioteca virtual, 63% les recomienda videos sobre temas académicos, 58,4% plantea cuestionarios o evaluaciones en la plataforma virtual, 47,3% contesta las preguntas de sus estudiantes por correo electrónico, 64,6% tiene una página web, blog o perfil de Facebook, 31,5% comenta en redes sociales y el 24,5% chatea eventualmente con los estudiantes sobre temas académicos.

En el **perfil de uso del Internet para el entretenimiento**, los usos más populares que se estudiaron son las redes sociales, chat, y videos de entretenimiento en YouTube, siendo la mayor actividad el uso de redes sociales, específicamente Facebook.

El número de horas que chatea por diversión es 14 en promedio, el número de horas a la semana que utiliza redes sociales son 16, el número de horas por semana que descarga música, videos y programas es de 7 horas, el número de videos para entretenimiento mira en YouTube cada semana es 10 videos por semana; en la apreciación de la utilidad del Smartphone con cámara fotográfica y acceso a Internet, teléfono móvil con acceso a Internet, computador portátil, Tablet (iPad; Galaxy Tab, Kindle, etc), cámara digital, iPod / MP3 Player, sobresale el computador portátil como dispositivo preferido por los estudiantes.

También se determinó diferencias significativas en relación a si tiene cuenta en YouTube, el número de personas con las que se maneja en redes sociales, y la confianza en la información de internet para realizar sus tareas.

5.3. Discusión de las hipótesis

Respecto de la Hipótesis 1, de los resultados de la investigación se obtiene que el nivel de ingresos económicos de los estudiantes, no incide en las actividades de uso de Internet para el aprendizaje, debido a que la mayor proporción de encuestados tiene acceso desde sus casas utilizando sus propios dispositivos, esto es a causa de la masificación del acceso al Internet y del abaratamiento de los costos de servicios, en relación con estudios anteriores no se concuerda, puesto que se indica que los estudiantes con niveles más altos tienen mayor acceso a Internet y lo usan con fines informativos y educativos como lo mencionan en Peter & Valkenburg (2006).

Al aplicar la regresión logística binaria para comprobar el modelo de la hipótesis 2, para los niveles de ingresos menores a 350 y 600 dólares se tiene un valor de significancia ($p < 0.05$), en los cuales los ingresos determinan el uso del Internet para actividades de entretenimiento, pero para ingresos mayores no hay significancia, por ello al interpretar los ODDs se obtuvo que la probabilidad de que el estudiante pertenezca al conglomerado de Bajo_Entretenimiento es 0.282 veces mayor en comparación con el nivel de 1500 dólares (OR=0.282 IC=95% 0.110 – 0.720 $p=0.008$). Las demás categorías no son significativas, por lo que se determina que no hay relación entre los ingresos y el uso de Internet para actividades de entretenimiento, estos estudios concuerdan con Zaragoza (2015), en el que indica que los estudiantes usan el Internet para fines de entretenimiento sin que exista relación con los ingresos económicos.

Para las hipótesis 3 y 4, no se encuentra relación entre el uso de la tecnología para el aprendizaje y el entretenimiento, con el rendimiento académico, esto debido a que los estudiantes se conectan desde sus casas, casi todos los días por varias horas para realizar actividades de entretenimiento y académicas, y sin embargo muestran un buen rendimiento académico, al aplicar las regresiones logísticas binarias no se encontró relación entre estas variables, por lo que no se puede deducir que mientras el estudiante invierta más tiempo en recursos educativos en línea mejorará su rendimiento académico, y tampoco podemos deducir que si el estudiante gasta su tiempo en contenido de entretenimiento esto afectara su rendimiento académico, estos datos concuerdan con Castaño (2011), que indica que el utilizar Internet para actividades académicas no se transforma en un mejor rendimiento académico, pero difiere del estudio de Castro (2013), quien indica que existe una relación entre el uso de redes sociales como actividades de entretenimiento y el rendimiento académico.

CONCLUSIONES

Se demostró que no hay relación entre el nivel de ingresos económicos y el uso del Internet para las actividades académicas de los estudiantes; considerando que el 57% tienen ingresos menores a 600 dólares mensuales, esto no afecta en el acceso a Internet desde los hogares que es del 71%, ni al número de horas que utilizan para conectarse al Internet con fines académicos.

Se determina que el nivel de ingresos económicos y el uso del Internet para actividades de entretenimiento tiene incidencia únicamente cuando los ingresos son menores a 600 dólares, y no tienen incidencia cuando los ingresos son mayores a 600 dólares, debido a que los estudiantes independientemente de los ingresos económicos tienen un alto acceso al Internet desde sus casas tanto para fines académicos como para entretenimiento, y puesto que al analizar los ODD, se concluye que el aumento del nivel de ingresos de las familias disminuye la probabilidad de la incidencia del Internet para el entretenimiento.

Mediante la comprobación de los modelos para las hipótesis 3 y 4 se determina que no existe relación entre las variables, debido a que al introducir las variables independientes (usos de la tecnología para el aprendizaje y usos de la tecnología para entretenimiento) sobre la variable dependiente (rendimiento académico) no son significativas para el modelo construidos con la regresión logística binomial.

Para este trabajo de titulación el uso de la herramienta SPSS fue de mucha utilidad por su versatilidad para manejar las tablas de datos en todas las etapas de la metodología KDD, además este software cuenta con los métodos y técnicas necesarias para realizar los diferentes procedimientos técnicos que necesarios para implementar los procedimientos y obtener información relevante referente a las variables de estudio.

La minería de datos ha experimentado un gran desarrollo relacionado con el crecimiento del volumen de datos que procesan los sistemas de información, mediante estas técnicas el investigador es capaz de obtener nuevas variables de estudio que no son perceptibles a simple vista pero que permiten obtener características especiales de la información. Sin embargo, estas técnicas requieren tener cuidado en el tratamiento de los datos, siendo la etapa de limpieza, fundamental para disminuir el nivel de error y lograr un mayor acercamiento a la realidad. La minería de datos, aun es un campo de estudio en desarrollo que requiere mayor investigación y optimización de las técnicas actualmente empleadas.

RECOMENDACIONES

Se recomienda que las instituciones educativas en todos los niveles, investiguen acerca de la incidencia de los usos de la tecnología, utilizando metodologías de minería de datos y análisis estadístico para descubrir conocimiento que brinde la posibilidad de implementar o mejorar metodologías de enseñanza aprendizaje; las mismas que permitan a futuros investigadores obtener datos relevantes del comportamiento de los estudiantes debido a que como se mencionó en la problemática de este documento, la sociedad está cambiando y seguirá cambiando, por lo tanto esta investigación debe ser continua pero su aplicación debe enmarcarse en la población de estudio, para formular políticas y acciones que fomenten el buen uso de Internet y motiven la investigación.

Considerando que el uso de la tecnología no incide en el rendimiento académico de los estudiantes de la ESPOCH, se recomienda implementar metodologías que apunten a mejorar las habilidades del aprendizaje, mediante la masificación del uso de herramientas e-learning y a través de una adecuada capacitación y orientación a los estudiantes y docentes quienes serán capaces de obtener mayores beneficios en el buen uso del Internet como una herramienta de apoyo para sus actividades académicas y así se logre una mejora en la en el uso de la tecnología para la educación virtual.

En cuanto a la minería de datos se recomienda al investigador que, en todas las etapas, encuentre los datos que le resulten extraños, trate de averiguar qué significan y si estos son valores lógicos con relación a la variable de estudio, para esto es importante buscar la fuente de los datos, puesto que la información valiosa es la que se obtiene al combinar las variables para hacerlas más significativas.

Para evitar errores en la investigación no se recomienda utilizar muestreos parciales, debido a que las variables pueden ser analizadas de forma dependiente o independiente, nunca nada debe suponerse, siempre debe desglosar la información para tener una visión precisa de la información.

TRABAJOS FUTUROS

Con los datos obtenidos en la encuesta aplicada, se puede recomendar trabajos futuros de estudio referentes a la incidencia del género en la carrera que estudia, y si el género influye o no en el rendimiento académico. Otro tema relevante a investigar, sería la relación que tienen el tiempo invertido al usar el Internet frente al tiempo invertido en actividades académicas.

Sería importante realizar un estudio similar al ejecutado, en donde se actualicen las preguntas de la encuesta en función del contexto social en que vivimos, en la pregunta de dispositivos que se utilizan, es importante considerar la masificación del acceso a planes de datos para teléfonos inteligentes. También es importante tomar en cuenta que la encuesta limita al investigador a hablar de tecnología únicamente refiriéndose al uso del Internet, por cuanto, no se indaga sobre otros equipos tecnológicos y los mecanismos de acceso a la red que se puede utilizar.

Actualmente, no solo los estudiantes universitarios acceden a la información en el Internet, el grupo más vulnerable son los niños, sobre todo los que sus padres han remplazado a la niñera por una Tablet, para que jueguen o miren videos, por ello se recomienda plantear como un trabajo futuro se realice esta investigación en las escuelas del país, para determinar cómo influye el acceso al Internet en las etapas del aprendizaje de los niños, y como el uso de la tecnología distrae a los niños y los aísla del mundo que les rodea.

REFERENCIAS

- Brunner, J. J., Brunner, J. J., & Tedesco, J. C. (2003). La educación al encuentro de las nuevas tecnologías. *Las nuevas tecnologías y el futuro de la educación*, 15–68.
- Camacho, K. (2005). La brecha digital. *Palabras en juego: enfoques multiculturales sobre las sociedades de la información*, 61–71.
- Castaño, C., & others. (2008). *La segunda brecha digital*. JSTOR. Recuperado a partir de <http://www.jstor.org/stable/pdf/40184925.pdf>
- Certad, C. C. (2010a, noviembre 25). 5 Programas Libres para Data Mining (Minería de Datos). Recuperado a partir de <http://fraterneo.blogspot.com/2010/11/5-programas-libres-para-data-mining.html>
- Certad, C. C. (2010b, noviembre 25). 5 Programas Libres para Data Mining (Minería de Datos). Recuperado a partir de <http://fraterneo.blogspot.com/2010/11/5-programas-libres-para-data-mining.html>
- Clustering - EcuRed. (2016, abril 16). Recuperado 24 de mayo de 2016, a partir de <http://www.ecured.cu/Clustering>
- Cobo, A. (s. f.). *Diseño y programación de bases de datos*. Madrid España: Editorial Visión Libros.
- Costas, G. (s. f.). Data Mining. Recuperado 13 de junio de 2016, a partir de <http://www.authorstream.com/Presentation/carmensanchez-391727-datamining601-education-ppt-powerpoint/>
- D3 Clustering.pdf. (s. f.). Recuperado a partir de <http://elvex.ugr.es/decsai/intelligent/slides/dm/D3%20Clustering.pdf>
- Fernández, J. M. B. (2003). *Gestión científica empresarial: temas de investigación actuales*. Netbiblo.
- Garcés H., D. R. (2011). *Mundos Virtuales*. Riobamba: (no publicado).
- Garre, M., Cuadrado, J. J., & Silicia, M. A. (s. f.). Comparación de diferentes algoritmos de clustering en la estimación de coste en el desarrollo de software, 11.

- Gonzales M., J. A. (2015). *UF0348 - Utilización de las bases de datos relacionales en el sistema de gestión y almacenamiento de datos*. Ediciones Paraninfo, S.A.
- IBM - SPSS Statistics Base. (2016, enero 1). Recuperado 23 de mayo de 2016, a partir de <http://www-03.ibm.com/software/products/es/spss-stats-base>
- image00118.png (844x101). (s. f.). Recuperado 24 de mayo de 2016, a partir de <https://unpocodejava.files.wordpress.com/2011/10/image00118.png>
- INEC. (2016, mayo 18). *infografia.pdf* [OFICIAL]. Recuperado 18 de mayo de 2016, a partir de http://www.inec.gob.ec/sitio_tics/infografia.pdf
- Información general sobre el procesamiento analítico en línea (OLAP) - Excel. (s. f.). Recuperado 23 de mayo de 2016, a partir de <https://support.office.com/es-es/article/Informaci%C3%B3n-general-sobre-el-procesamiento-anal%C3%ADtico-en-l%C3%ADnea-OLAP-15d2cdde-f70b-4277-b009-ed732b75fdd6>
- Izarra, C. (2010, octubre). Artículo: Mobile Learning. Recuperado a partir de <https://carolinaizarra.wordpress.com/81-2/>
- KDD Proceso de Extracción de conocimiento | WebMining. (2011). Recuperado 24 de mayo de 2016, a partir de <http://www.webmining.cl/2011/01/proceso-de-extraccion-de-conocimiento/>
- Krall, C. (2006). minería de datos. Recuperado 22 de mayo de 2016, a partir de http://www.aprenderaprogramar.com/index.php?option=com_attachments&task=download&id=203
- La internet.... (2009, julio 27). Recuperado a partir de <https://aleja11a.wordpress.com/2009/07/27/la-internet/>
- López, R. M. (2011). *Mundos virtuales 3D: Una guía para padres y formadores*. Editorial UOC.
- Martinez, E., & Salanova, S. (s. f.). La evaluación de los aprendizaje. Recuperado 20 de mayo de 2016, a partir de <http://www.uhu.es/cine.educacion/didactica/0091evaluacionaprendizaje.htm>
- Medina, P. (2011). REDES SOCIALES EN LA EDUCACIÓN [EDUCACION]. Recuperado 18

- de mayo de 2016, a partir de <http://pucesa.edu.ec/content/redes-sociales-en-la-educaci%C3%B3n>
- Mi página de Weka. (s. f.). Recuperado 24 de mayo de 2016, a partir de <http://users.dsic.upv.es/~cferri/weka/>
- Microsoft SQL Server. (2016). Recuperado 23 de mayo de 2016, a partir de [https://technet.microsoft.com/es-es/library/mt590198\(v=sql.1\).aspx](https://technet.microsoft.com/es-es/library/mt590198(v=sql.1).aspx)
- Molinero, L. (2002). Análisis de la covarianza. Recuperado 23 de mayo de 2016, a partir de <http://www.seh-lelha.org/ancova.htm>
- Moreira, M. A., & Segura, J. A. (2009). E-learning: enseñar y aprender en espacios virtuales. *Tecnología educativa. La formación del profesorado en la era de Internet, Málaga, Aljibe*. Recuperado a partir de https://www.researchgate.net/profile/Manuel_Area/publication/216393113_E-Learning_ensenar_y_aprender_en_espacios_virtuales/links/0c96051ebd02aca3660000.pdf
- Ortiz, P. (2015, noviembre 7). 6 herramientas gratuitas para análisis de datos. Recuperado a partir de <http://mprende.co/gesti%C3%B3n/6-herramientas-gratuitas-para-an%C3%A1lisis-de-datos>
- Pérez López, C., & Santín Gonzáles, D. (2007a). *Minería de datos: técnicas y herramientas*. Editorial Paraninfo.
- Pérez López, C., & Santín Gonzáles, D. (2007b). *Minería de datos: técnicas y herramientas*. Editorial Paraninfo.
- RapidMiner | #1 Open Source Predictive Analytics Platform. (2016, mayo 24). Recuperado 24 de mayo de 2016, a partir de <https://rapidminer.com/>
- RapidMiner | #1 Open Source Predictive Analytics Platform. (s. f.). Recuperado 24 de mayo de 2016, a partir de <https://rapidminer.com/>
- Sánchez, J. (s. f.). Internet y el World Wide Web. Recuperado 22 de mayo de 2016, a partir de http://ific.uv.es/wop/SABER_MAS/internet.html

- Santoyo, A. S., & Martínez, E. M. (2003). *La brecha digital: mitos y realidades*. UABC.
- Severin, E. (2013). enfoques estrategicos sobre las tics en educaicon en america latina y el caribe. Recuperado 19 de mayo de 2016, a partir de <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/images/ticsesp.pdf>
- Soline Moreno, J. C. (2016, revisado demayo). calidad-costos-accesos-servicios-internet-ecuador.pdf. Recuperado 18 de mayo de 2016, a partir de <http://www.infodesarrollo.ec/documentos/items/show/29>.
- SPSS. (2016, mayo 23). En *Wikipedia, la enciclopedia libre*. Recuperado a partir de <https://es.wikipedia.org/w/index.php?title=SPSS&oldid=91241049>
- Torres, J. C. (2012). Análisis de las relaciones entre los niveles de ingreso, edad y género de los estudiantes, los usos de internet y el rendimiento académico en un grupo de universidades ecuatorianas presenciales. Universitat Obeerta de Catalunya.
- UNESCO. (2016). Las TIC en la educación | Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [OFICIAL]. Recuperado 18 de mayo de 2016, a partir de <http://www.unesco.org/new/es/unesco/themes/icts/>
- Viera, L., Ortiz, L., & Ramirez, S. (2009). *Introducción a la Minería de Datos*. Editora E-papers.
- Zapata, B., & Rojas, H. (2014). La brecha digital universitaria: La apropiación de las TIC en estudiantes de educación superior en Bogotá (Colombia). *43, XXII 2014*, 133-142. <http://doi.org/http://dx.doi.org/10.3916/C43-2014-13>
- Zaragozión, A. (2016). *Analisis de los usos de la tecnologia en los estudiantes de los colegios de la ciudad de Cuenca*.

ANEXOS

Anexo 1: Entrevista

Guía para la entrevista a un(a) directivo de la Institución Educativa

1. Instalaciones físicas

- ¿Cuenta con salas de cómputo para los estudiantes?
- ¿Cuenta con Internet inalámbrico para los estudiantes?
- ¿Ancho de banda total?
- ¿Número de estudiantes?
- ¿Tiene software para la gestión académica y administrativa?
Desarrollo propio _____
Sistema comercial _____
- ¿Cuenta con un campus virtual (entorno virtual de aprendizaje)?
Desarrollo propio: _____
Sistema comercial: _____ Nombre: _____
No tiene _____
- ¿Los estudiantes se pueden matricular en línea?
- ¿Los estudiantes pueden consultar sus calificaciones en línea?
- ¿Los estudiantes y profesores tienen acceso a una biblioteca virtual?
- ¿Los profesores suben las calificaciones a Internet?
- ¿Qué porcentaje de profesores cuenta con un computador personal?

2. Políticas de uso de tecnología

- ¿Los docentes utilizan obligatoriamente elementos tecnológicos en sus actividades de enseñanza? Si _____ NO _____
- ¿Cuenta la institución con un plan de formación de profesores en el uso de tecnologías para la educación?
- ¿Aproximadamente qué porcentaje de sus profesores está capacitado en temas tecnológicos?
- ¿Aproximadamente qué porcentaje de sus profesores está capacitado en temas pedagógicos?
- ¿Tienen algún tipo de incentivo para los profesores que utilizan las TICs en sus actividades académicas?

Anexo 2: Encuesta

Estimado estudiante, solicitamos su colaboración contestando esta encuesta, la que permitirá desarrollar una investigación para conocer el uso de Internet en la Escuela Superior Politécnica de Chimborazo.

1. Responda la siguiente pregunta	
¿En qué universidad estudia?	

2. Responda la siguiente pregunta	
¿Qué carrera estudia?	

3. Responda la siguiente pregunta	
¿Cuál es su edad?	

4. Responda la siguiente pregunta	Hombre	Mujer
¿Cuál es su género?	()	()

5. Los ingresos mensuales de su familia son de:	
Hasta 350 dólares	()
Hasta 600 dólares	()
Hasta 1.000 dólares	()
Hasta 1.500 dólares	()
Más de 1.500 dólares	()

6. ¿Desde dónde se conecta habitualmente a Internet? (escoja solo una opción)	
Desde la casa	()
Desde un cyber café	()
Desde el trabajo	()
Desde la Universidad	()
Desde una red móvil (movistar, claro, cnt)	()

7. Responda la siguiente pregunta	1	2	3	4	5	6	7
De 1 a 7, ¿cuántos días a la semana se conecta Internet?	()	()	()	()	()	()	()

8. Responda las siguientes preguntas	1	2	3	4	5	6	7	8	9	10
De 1 a 10 su nivel de conocimientos en el manejo de Internet es:	()	()	()	()	()	()	()	()	()	()

9. Responda las siguientes preguntas	
¿Aproximadamente cuántas horas se conecta cada día?	(____)
¿Hace cuántos años se conecta a Internet?	(____)

10. En lo referente a las asignaturas en las que está matriculado	
¿Cuántas veces a la semana ingresa a la plataforma virtual de su universidad?	(____)
¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	(____)
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	(____)
¿Aproximadamente cuántos recursos educativos descarga de la plataforma virtual cada mes?	(____)
¿Aproximadamente cuántos videos académicos mira en youtube cada mes?	(____)
¿Aproximadamente en cuántos foros virtuales participa cada mes?	(____)
¿Aproximadamente cuántos post o tweets sobre temas académicos realiza en las redes sociales por mes?	(____)
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	(____)
¿Aproximadamente cuántas horas busca información académica en internet cada mes?	(____)
¿Aproximadamente cuántas horas utiliza la biblioteca virtual de la universidad cada mes?	(____)

11. En lo referente al entretenimiento y diversión en internet	
¿Aproximadamente cuántas horas a la semana chatea por diversión?	(____)
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	(____)
¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	(____)

¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	(____)
¿Aproximadamente cuántos videos para entretenimiento mira en youtube cada semana?	(____)

12. Responda las siguientes preguntas	
¿ Aproximadamente cuántos seguidores tiene en twitter?	(____)
¿ Cuántos amigos tiene en facebook?	(____)
¿ Cuántos contactos tiene en LinkedIn?	(____)

13. Responda con una X en SI o NO a las siguientes preguntas	SI	No
Tiene un blog	()	()
Tiene cuenta en youtube	()	()
Tiene cuenta en www.del.icio.us	()	()
	()	()

14. ¿Cuál es su nivel de uso de los siguientes dispositivos? (1 significa no usar y 10 significa utilizar al máximo)	1	2	3	4	5	6	7	8	9	10
Smartphone con cámara fotográfica y acceso a internet	0	0	0	0	0	0	0	0	0	0
Teléfono móvil con acceso a internet	0	0	0	0	0	0	0	0	0	0
Teléfono móvil sin acceso a internet	0	0	0	0	0	0	0	0	0	0
Computador portátil	0	0	0	0	0	0	0	0	0	0
Tablet (iPad; Galaxy Tab, Kindle, etc)	0	0	0	0	0	0	0	0	0	0
Cámara digital	0	0	0	0	0	0	0	0	0	0
iPod / MP3 Player	0	0	0	0	0	0	0	0	0	0

15. De 1 a 10 valore los siguientes aspectos (1 significa no estar de acuerdo y 10 estar completamente de acuerdo)	1	2	3	4	5	6	7	8	9	10
Internet le permite elaborar los trabajos más rápido y con menos esfuerzo	0	0	0	0	0	0	0	0	0	0
Usted confía en la información de internet para realizar sus tareas	0	0	0	0	0	0	0	0	0	0
Internet le permite prescindir de la Biblioteca	0	0	0	0	0	0	0	0	0	0
Internet facilita el proceso de aprendizaje	0	0	0	0	0	0	0	0	0	0
Internet le permite mejorar sus calificaciones	0	0	0	0	0	0	0	0	0	0
Usted presenta trabajos académicos copiados desde Internet	0	0	0	0	0	0	0	0	0	0

16. Responda las siguientes preguntas referentes a sus profesores. (Se recomienda evaluar de forma general a todos sus profesores)	SI	NO	A veces
Su profesor ingresa a la plataforma virtual	()	()	()
Contesta sus consultas por correo electrónico	()	()	()
Chatea con usted eventualmente sobre aspectos académicos	()	()	()
Su profesor comenta en redes sociales sobre temas académicos	()	()	()
Le sube materiales digitales a la plataforma virtual	()	()	()
Le recomienda recursos digitales de la biblioteca virtual	()	()	()
Le recomienda videos sobre temas académicos	()	()	()
Le plantea cuestionarios o evaluaciones en la plataforma virtual	()	()	()
Le plantea foros virtuales	()	()	()
Su profesor tiene una página web, blog o perfil de facebook	()	()	()
Su profesor tiene cuenta de twitter	()	()	()

17. Responda las siguientes preguntas:	
En el semestre anterior, ¿en cuántas asignaturas se matriculó?	()
En el semestre anterior ¿cuántas asignaturas aprobó?	()

Anexo 3: Estadísticos descriptivos en SPSS

Pregunta 1

¿En qué universidad estudia?	
ESCUELA SUPERIOR POLITECNICA DE CHIMBORAZO	392

Pregunta 2

¿Que carrera estudia?		
	¿Que carrera estudia?	¿Que carrera estudia?
Válidos	FADE - INGENIERÍA EN CONTABILIDAD Y AUDITORÍA	51
	FADE - INGENIERÍA EN EMPRESAS	1
	FC - BIOQUÍMICA Y FARMACIA	16
	FC - INGENIERÍA EN BIOFÍSICA	25
	FC - INGENIERÍA EN BIOTECNOLOGÍA AMBIENTAL	8
	FC - INGENIERÍA EN ESTADÍSTICA INFORMÁTICA	4
	FC - INGENIERÍA QUÍMICA	3
	FCP - INGENIERÍA EN INDUSTRIAS PECUARIAS	18
	FIE - INGENIERÍA EN ELECTRÓNICA, CONTROL Y REDES INDUSTRIALES	13
	FIE - INGENIERÍA EN ELECTRÓNICA, TELECOMUNICACIONES Y REDES	42
	FIE - INGENIERÍA EN SISTEMAS INFORMÁTICOS	15
	FM - INGENIERÍA AUTOMOTRIZ	21
	FM - INGENIERÍA DE MANTENIMIENTO	3
	FM - INGENIERÍA INDUSTRIAL	44
	FM - INGENIERÍA MECÁNICA	14
	FRN - INGENIERÍA EN ECOTURISMO	1
	FRN - INGENIERÍA ZOOTÉCNICA	19
	FSP - GASTRONOMÍA	7
	FSP - MEDICINA	39
	FSP - NUTRICIÓN Y DIETÉTICA	39
	FSP - PROMOCIÓN Y CUIDADOS DE LA SALUD	9
	Total	392

Pregunta 3

¿Cuál es su edad?					
		¿Cuál es su edad?	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	18	26	6,6	6,6	6,6
	19	52	13,3	13,3	19,9
	20	82	20,9	20,9	40,8
	21	52	13,3	13,3	54,1
	22	59	15,1	15,1	69,1
	23	56	14,3	14,3	83,4
	24	32	8,2	8,2	91,6
	25	17	4,3	4,3	95,9
	26	11	2,8	2,8	98,7
	27	3	,8	,8	99,5
	29	1	,3	,3	99,7
	32	1	,3	,3	100,0
		Total	392	100,0	100,0

Pregunta 4

¿Cuál es su género?					
		Genero	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Hombre	183	46,7	46,7	46,7
	Mujer	209	53,3	53,3	100,0
	Total	392	100,0	100,0	

Pregunta 5

Los ingresos mensuales de su familia son de:					
		Ingersos Mensuales	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Hasta 350 dólares	89	22,7	22,7	22,7
	Hasta 600 dólares	133	33,9	33,9	56,6
	Hasta 1000 dólares	86	21,9	21,9	78,6
	Hasta 1.500 dólares	48	12,2	12,2	90,8
	Más de 1.500 dólares	36	9,2	9,2	100,0
	Total	392	100,0	100,0	

Pregunta 6

¿Desde dónde se conecta habitualmente a Internet? (escoja solo una opción)					
		Lugar de Conexión a Internet	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Desde la casa	277	70,7	70,7	70,7
	Desde un cyber café	55	14,0	14,0	84,7
	Desde el trabajo	2	,5	,5	85,2
	Desde la Universidad	30	7,7	7,7	92,9
	Desde una red móvil (movistar, claro, cnt)	28	7,1	7,1	100,0
	Total	392	100,0	100,0	

Pregunta 7

De 1 a 7, ¿cuántos días a la semana se conecta Internet?					
		Días de Conexión a Internet	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1	3	,8	,8	,8
	2	4	1,0	1,0	1,8
	3	16	4,1	4,1	5,9
	4	31	7,9	7,9	13,8
	5	59	15,1	15,1	28,8
	6	62	15,8	15,8	44,6
	7	217	55,4	55,4	100,0
	Total	392	100,0	100,0	

Pregunta 8

De 1 a 10 su nivel de conocimientos en el manejo de Internet es:					
		Manejo de Internet	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	2	3	,8	,8	,8
	3	5	1,3	1,3	2,0
	4	15	3,8	3,8	5,9
	5	40	10,2	10,2	16,1
	6	50	12,8	12,8	28,8
	7	81	20,7	20,7	49,5
	8	134	34,2	34,2	83,7
	9	38	9,7	9,7	93,4
	10	26	6,6	6,6	100,0
	Total	392	100,0	100,0	

Estadísticos		
De 1 a 10 su nivel de conocimientos en el		
N	Válidos	392
	Perdidos	0
Media		7,1990
Mediana		8,0000
Desv. típ.		1,60281
Mínimo		2,00
Máximo		10,00

Pregunta 9

Estadísticos			
		¿Aproximadamente cuántas horas se conecta cada día?	¿Hace cuántos años se conecta a Internet?
N	Válidos	392	392
	Perdidos	0	0
Media		5,6454	6,4821
Mediana		5,0000	6,0000
Moda		5,00	5,00
Desv. típ.		4,15361	3,37354
Mínimo		1,00	1,00
Máximo		24,00	23,00
Suma		2213,00	2541,00

Pregunta 10

Estadísticos											
		¿Cuántas veces a la semana ingresa a la plataforma virtual de su universidad?	¿Aproximadamente cuántas consultas les hace a sus profesores cada mes?	¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	¿Aproximadamente cuántos recursos educativos descarga de la plataforma virtual cada mes?	¿Aproximadamente cuántos videos académicos mira en youtube cada mes?	¿Aproximadamente en cuántos foros virtuales participa cada mes?	¿Aproximadamente cuántos post o tweets sobre temas académicos realiza en las redes sociales por mes?	¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	¿Aproximadamente cuántas horas busca información académica en internet cada mes?	¿Aproximadamente cuántas horas utiliza la biblioteca virtual de la universidad cada mes?
N	Válidos	392	392	392	392	392	392	392	392	392	392
	Perdidos	0	0	0	0	0	0	0	0	0	0
Media		4	6	9	6	8	2	4	15	22	5
Mediana		3	4	6	3	5	1	1	5	10	2
Moda		3	2	10	2	2	0	0	2	20	0
Mínimo		0	0	0	0	0	0	0	0	0	0
Máximo		51	50	100	150	200	40	100	1050	160	150

Pregunta 11

Estadísticos						
		¿Aproximadamente cuántas horas a la semana chatea por diversión?	¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	¿Aproximadamente cuántas horas a la semana utiliza juegos en línea?	¿Aproximadamente cuántas horas a la semana descarga música, videos y programas?	¿Aproximadamente cuántos videos para entretenimiento mira en youtube cada semana?
		Chat	Redes Sociales	Juegos en Linea	Descargas	YouTube
Media		14,77	16,14	4,24	7,09	10,89
Moda		2,00	10,00	0,00	1,00	2,00
Mínimo		0,00	0,00	0,00	0,00	0,00
Máximo		60,00	70,00	40,00	50,00	50,00

Pregunta 12

Estadísticos				
		¿Aproximadamente cuántos seguidores tiene en twitter?	¿Cuántos amigos tiene en facebook?	¿Cuántos contactos tiene en LinkedIn?
		Twitter	Facebook	LinkedIn
Media		47,2015	519,1964	10,7194
Moda		0,00	300,00	0,00
Mínimo		0,00	0,00	0,00
Máximo		3984,00	8000,00	1784,00

Pregunta 13

Estadísticos				
		Tiene un blog	Tiene cuenta en youtube	Tiene cuenta en www.del.icio.us
N	Válidos	392	392	390
	Perdidos	0	0	2
Media		1,5969	1,3036	1,9462
Mediana		2,0000	1,0000	2,0000
Moda		2,00	1,00	2,00
Mínimo		1,00	1,00	1,00
Máximo		2,00	2,00	2,00

Pregunta 14

		Estadísticos						
		Smartphone con cámara fotográfica y acceso a internet	Teléfono móvil con acceso a internet	Teléfono móvil sin acceso a internet	Computador portátil	Tablet (iPad; Galaxy Tab, Kindle, etc)	Cámara digital	iPod / MP3 Player
N	Válidos	392	392	392	392	392	392	392
	Perdidos	0	0	0	0	0	0	0
Media		5,2041	5,8087	5,2730	7,1250	4,1429	4,5791	4,0077
Moda		1,00	1,00	1,00	10,00	1,00	1,00	1,00
Mínimo		1,00	1,00	1,00	1,00	1,00	1,00	1,00
Máximo		10,00	10,00	10,00	10,00	10,00	10,00	10,00

Pregunta 15

		Estadísticos					
		Internet le permite elaborar los trabajos más rápido y con menos esfuerzo	Usted confía en la información de internet para realizar sus tareas	Internet le permite prescindir de la Biblioteca	Internet facilita el proceso de aprendizaje	Internet le permite mejorar sus calificaciones	Usted presenta trabajos académicos copiados desde Internet
N	Válidos	392	392	392	392	392	392
	Perdidos	0	0	0	0	0	0
Media		7,7526	6,1429	6,6020	7,4337	6,7347	5,9107
Moda		10,00	6,00	8,00	10,00	8,00	8,00
Mínimo		1,00	1,00	1,00	1,00	1,00	1,00
Máximo		10,00	10,00	10,00	10,00	10,00	10,00

Pregunta 16

		Estadísticos										
		Su profesor ingresa a la plataforma virtual	Contesta sus consultas por correo electrónico	Chatea con usted eventualmente sobre aspectos académicos	Su profesor comenta en redes sociales sobre temas académicos	Le sube materiales digitales a la plataforma virtual	Le recomienda recursos digitales de la biblioteca virtual	Le recomienda videos sobre temas académicos	Le plantea cuestionarios o evaluaciones en la plataforma virtual	Le plantea foros virtuales	Su profesor tiene una página web, blog o perfil de facebook	Su profesor tiene cuenta de twitter
N	Válidos	392	392	392	392	392	392	390	392	392	392	392
	Perdidos	0	0	0	0	0	0	2	0	0	0	0
Media		1,5281	1,7781	1,9745	1,8648	1,3648	1,4949	1,5538	1,6199	1,6276	1,4541	1,8469
Moda		1,00	1,00	2,00	2,00	1,00	1,00	1,00	1,00	1,00	1,00	2,00
Mínimo		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Máximo		3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00

Pregunta 17

Estadísticos				
		En el semestre anterior, ¿en cuántas asignaturas se matriculó?	En el semestre anterior ¿cuántas asignaturas aprobó?	Materias Reprobadas V56-V57
N	Válidos	392	392	392
	Perdidos	0	0	0
Media		6,5102	6,1990	,3138
Moda		8,00	6,00	0,00
Mínimo		1,00	0,00	0,00
Máximo		10,00	10,00	5,00

En el semestre anterior ¿cuántas asignaturas aprobó?					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	,00	1	,3	,3	,3
	2,00	9	2,3	2,3	2,6
	3,00	23	5,9	5,9	8,4
	4,00	34	8,7	8,7	17,1
	5,00	47	12,0	12,0	29,1
	6,00	98	25,0	25,0	54,1
	7,00	72	18,4	18,4	72,4
	8,00	95	24,2	24,2	96,7
	9,00	10	2,6	2,6	99,2
	10,00	3	,8	,8	100,0
	Total	392	100,0	100,0	

Materias Reprobadas V56-V57					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	,00	305	77,8	77,8	77,8
	1,00	62	15,8	15,8	93,6
	2,00	19	4,8	4,8	98,5
	3,00	2	,5	,5	99,0
	4,00	3	,8	,8	99,7
	5,00	1	,3	,3	100,0
	Total	392	100,0	100,0	

Anexo 4: K-Medias para el uso de Internet en actividades académicas

- Dos grupos

Centros de los conglomerados finales		
	Conglomerado	
	Bajo_Aprendizaje (10%)	Alto_Aprendizaje (90%)
¿Aproximadamente cuántas horas busca información académica en internet cada mes?	14,63	90,45
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	8,06	46,73
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	8,09	19,35
¿Aproximadamente cuántos videos académicos mira en youtube cada mes?	6,47	23,45

ANOVA						
	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas busca información académica en internet cada mes?	206510,161	1	237,109	390	870,951	,000
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	53698,268	1	242,015	390	221,880	,000
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	4555,566	1	111,793	390	40,750	,000
¿Aproximadamente cuántos videos académicos mira en youtube cada mes?	10357,523	1	200,332	390	51,702	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado			
Conglomerado	Alto_Aprendizaje	352,000	90%
	Bajo_Aprendizaje	40,000	10%
Válidos		392,000	100%
Perdidos		0,000	

- **Tres grupos**

Centros de los conglomerados finales			
	Conglomerado		
	1	2	3
¿Aproximadamente cuántas horas busca información académica en internet cada mes?	14,42	100,00	87,32
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	7,93	100,00	40,24
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	8,07	34,00	17,39
¿Aproximadamente cuántos videos académicos mira en youtube cada mes?	6,43	125,00	12,18

ANOVA						
	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas busca información académica en internet cada mes?	103246,467	2	237,763	389	434,242	,000
¿Aproximadamente cuántas horas chatea sobre temas académicos cada mes?	33529,816	2	208,289	389	160,978	,000
¿Aproximadamente cuántas consultas realiza a sus compañeros por mes?	2729,751	2	109,757	389	24,871	,000
¿Aproximadamente cuántos videos académicos mira en youtube cada mes?	28131,690	2	82,837	389	339,602	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado	
Conglomerado	1
	2
	3
Válidos	392,000
Perdidos	0,000

- Cuatro grupos

Centros de los conglomerados finales				
	Conglomerado			
	1	2	3	4
¿Aproximadamente cuántas horas busca información académica en internet cada	14,20	66,00	93,03	100,00
¿Aproximadamente cuántas horas chatea sobre temas	7,93	91,36	18,62	100,00
¿Aproximadamente cuántas consultas realiza a sus	8,15	17,36	15,79	34,00
¿Aproximadamente cuántos videos académicos mira en	6,38	17,82	10,28	125,00

ANOVA						
	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas busca información académica en internet cada	71025,892	3	221,404	388	320,797	,000
¿Aproximadamente cuántas horas chatea sobre temas	35761,712	3	105,152	388	340,096	,000
¿Aproximadamente cuántas consultas realiza a sus	1612,467	3	111,643	388	14,443	,000
¿Aproximadamente cuántos videos académicos mira en	18953,827	3	81,509	388	232,536	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado	
Conglomerado	1
	2
	3
	4
Válidos	392,000
Perdidos	0,000

- **Análisis de discriminante**

Resultados de la clasificación ^a					
Número inicial de casos			pertenencia		Total
			1	2	
Original	Recuento	1	338	14	352
		2	0	40	40
	%	1	96,0	4,0	100,0
		2	0,0	100,0	100,0

a. Clasificados correctamente el 96,4% de los casos

Resultados de la clasificación ^a						
Número inicial de casos			pronosticado			Total
			1	2	3	
Original	Recuento	1	328	0	22	350
		2	0	4	0	4
		3	0	1	37	38
	%	1	93,7	0,0	6,3	100,0
		2	0,0	100,0	0,0	100,0
		3	0,0	2,6	97,4	100,0

a. Clasificados correctamente el 94,1% de los casos agrupados

Resultados de la clasificación ^a							
Número inicial de casos			Grupo de pertenencia pronosticado				Total
			1	2	3	4	
Original	Recuento	1	326	11	11	0	348
		2	0	10	0	1	11
		3	0	0	29	0	29
		4	0	0	0	4	4
	%	1	93,7	3,2	3,2	0,0	100,0
		2	0,0	90,9	0,0	9,1	100,0
		3	0,0	0,0	100,0	0,0	100,0
		4	0,0	0,0	0,0	100,0	100,0

a. Clasificados correctamente el 94,1% de los casos agrupados originales.

Anexo 5: K-Medias para el uso de Internet en actividades de entretenimiento

- Dos grupos

Centros de los conglomerados finales		
	Conglomerado	
	Bajo Entretenimiento (22%)	Alto Entretenimiento (78%)
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	9,22	41,14
¿Aproximadamente cuántas horas a la semana chatea por diversión?	8,17	38,62
¿Aproximadamente cuántos videos para entretenimiento mira en youtube cada semana?	7,72	22,34

ANOVA						
	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	67824,756	1	109,444	390	619,720	,000
¿Aproximadamente cuántas horas a la semana chatea por diversión?	61752,856	1	90,529	390	682,130	,000
¿Aproximadamente cuántos videos para entretenimiento mira en youtube cada semana?	14231,268	1	119,449	390	119,141	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado		
Conglomerado	Alto Entretenimiento	307,000
	Bajo Entretenimiento	85,000
Válidos		392,000
Perdidos		0,000

- Tres grupos

Centros de los conglomerados finales			
	Conglomerado		
	Alto	Bajo	Medio
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	47,18	9,21	26,62
¿Aproximadamente cuántas horas a la semana chatea por diversión?	39,23	8,01	34,50
¿Aproximadamente cuántos videos para entretenimiento mira en youtube cada semana?	11,82	7,13	42,79

ANOVA						
	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la semana utiliza redes sociales?	36095,949	2	98,499	389	366,460	,000
¿Aproximadamente cuántas horas a la semana chatea por diversión?	30266,454	2	93,898	389	322,332	,000
¿Aproximadamente cuántos videos para entretenimiento mira en youtube cada semana?	19467,646	2	56,249	389	346,096	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son iguales.

Número de casos en cada conglomerado		
Conglomerado	Alto	56,000
	Bajo	302,000
	Medio	34,000
Válidos		392,000
Perdidos		0,000

- Cuatro grupos

Centros de los conglomerados finales				
	Conglomerado			
	1	2	3	4
¿Aproximadamente cuántas horas a la	25,81	6,83	22,40	55,84
¿Aproximadamente cuántas horas a la	34,45	5,44	22,95	42,55
¿Aproximadamente cuántos videos para	43,71	6,53	10,27	13,58

ANOVA						
	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
¿Aproximadamente cuántas horas a la	29040,474	3	60,275	388	481,802	,000
¿Aproximadamente cuántas horas a la	22635,962	3	75,133	388	301,280	,000
¿Aproximadamente cuántos videos para	12788,707	3	57,861	388	221,024	,000

Las pruebas F sólo se deben utilizar con una finalidad descriptiva puesto que los conglomerados han sido elegidos para maximizar las diferencias entre los casos en diferentes conglomerados. Los niveles críticos no son corregidos, por lo que no pueden interpretarse como pruebas de la hipótesis de que los centros de los conglomerados son

Número de casos en cada conglomerado		
Conglomerado	1	31,000
	2	246,000
	3	77,000
	4	38,000
Válidos		392,000
Perdidos		0,000

- **Análisis de discriminante**

Resultados de la clasificación ^a					
Número inicial de casos			pertenencia		Total
			1	2	
Original	Recuento	1	290	17	307
		2	0	85	85
	%	1	94,5	5,5	100,0
		2	0,0	100,0	100,0

a. Clasificados correctamente el 95,7% de los casos agrupados originales.

Resultados de la clasificación ^a						
Número inicial de casos			pronosticado			Total
			1	2	3	
Original	Recuento	1	55	0	1	56
		2	12	281	9	302
		3	1	0	33	34
	%	1	98,2	0,0	1,8	100,0
		2	4,0	93,0	3,0	100,0
		3	2,9	0,0	97,1	100,0

a. Clasificados correctamente el 94,1% de los casos agrupados originales.

Resultados de la clasificación ^a							
Número inicial de casos			pronosticado				Total
			1	2	3	4	
Original	Recuento	1	31	0	0	0	31
		2	2	226	18	0	246
		3	3	0	70	4	77
		4	1	0	0	37	38
	%	1	100,0	0,0	0,0	0,0	100,0
		2	,8	91,9	7,3	0,0	100,0
		3	3,9	0,0	90,9	5,2	100,0
		4	2,6	0,0	0,0	97,4	100,0

a. Clasificados correctamente el 92,9% de los casos agrupados originales.

Anexo 6: Pruebas chi-cuadrado y medidas simétricas

Relación entre ingreso y uso de Internet para el aprendizaje

Pruebas de chi-cuadrado			
	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	3,789 ^a	4	,435
Razón de verosimilitudes	3,988	4	,408
Asociación lineal por lineal	,042	1	,837
N de casos válidos	392		

a. 2 casillas (20,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 3,67.

Medidas simétricas					
		Valor	Error típ. asint. ^a	T aproximada ^b	Sig. aproximada
Ordinal por ordinal	Tau-c de Kendall	,003	,030	,106	,915
	Correlación de Spearman	,005	,044	,093	,926 ^c
Intervalo por intervalo	R de Pearson	-,010	,043	-,205	,837 ^c
N de casos válidos		392			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

c. Basada en la aproximación normal.

Relación entre ingresos y uso de Internet para el entretenimiento

Pruebas de chi-cuadrado			
	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	12,014 ^a	4	,017
Razón de verosimilitudes	12,170	4	,016
Asociación lineal por lineal	11,155	1	,001
N de casos válidos	392		

a. 0 casillas (0,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 7,81.

Medidas simétricas					
		Valor	Error típ. asint. ^a	T aproximada ^b	Sig. aproximada
Ordinal por ordinal	Tau-c de Kendall	,159	,046	3,469	,001
	Correlación de Spearman	,173	,049	3,466	,001 ^c
Intervalo por intervalo	R de Pearson	,169	,050	3,384	,001 ^c
N de casos válidos		392			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

c. Basada en la aproximación normal.

Relación entre uso de Internet para el aprendizaje y rendimiento académico

Pruebas de chi-cuadrado					
	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	,018 ^a	1	,893		
Corrección por continuidad ^b	0,000	1	1,000		
Razón de verosimilitudes	,018	1	,893		
Estadístico exacto de Fisher				,856	,511
Asociación lineal por lineal	,018	1	,893		
N de casos válidos	392				

a. 0 casillas (0,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 11,63.

b. Calculado sólo para una tabla de 2x2.

Medidas simétricas					
		Valor	Error típ. asint. ^a	T aproximada ^b	Sig. aproximada
Ordinal por ordinal	Tau-c de Kendall	-,004	,028	-,134	,894
	Correlación de Spearman	-,007	,051	-,135	,893 ^c
Intervalo por intervalo	R de Pearson	-,007	,051	-,135	,893 ^c
N de casos válidos		392			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

c. Basada en la aproximación normal.

Relación entre uso de Internet para el entretenimiento y rendimiento académico

Pruebas de chi-cuadrado					
	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	,006 ^a	1	,940		
Corrección por continuidad ^b	0,000	1	1,000		
Razón de verosimilitudes	,006	1	,940		
Estadístico exacto de Fisher				1,000	,519
Asociación lineal por lineal	,006	1	,940		
N de casos válidos	392				

a. 0 casillas (0,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 24,72.

b. Calculado sólo para una tabla de 2x2.

Medidas simétricas					
		Valor	Error típ. asint. ^a	T aproximada ^b	Sig. aproximada
Ordinal por ordinal	Tau-c de Kendall	-,003	,038	-,076	,940
	Correlación de Spearman	-,004	,051	-,076	,940 ^c
Intervalo por intervalo	R de Pearson	-,004	,051	-,076	,940 ^c
N de casos válidos		392			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

c. Basada en la aproximación normal.

Anexo 7: Regresión Binomial Hipótesis 1

Resumen del procesamiento de los casos

		N	Porcentaje
Casos no ponderados ^a			
Casos seleccionados	Incluidos en el análisis	392	100,0
	Casos perdidos	0	0,0
	Total	392	100,0
Casos no seleccionados		0	0,0
Total		392	100,0

a. Si está activada la ponderación, consulte la tabla de clasificación para ver el número total de casos.

Codificación de la variable dependiente

Valor original	Valor interno
1	0
2	1

Codificaciones de variables categóricas

		Frecuencia	Codificación de parámetros			
			(1)	(2)	(3)	(4)
Los ingresos mensuales de su familia son de:	Hasta 350 dólares	89	1,000	0,000	0,000	0,000
	Hasta 600 dólares	133	0,000	1,000	0,000	0,000
	Hasta 1000 dólares	86	0,000	0,000	1,000	0,000
	Hasta 1.500 dólares	48	0,000	0,000	0,000	1,000
	Más de 1.500 dólares	36	0,000	0,000	0,000	0,000

Bloque 0: Bloque inicial

Tabla de clasificación^{a,b}

Observado		Pronosticado		
		Número inicial de casos		Porcentaje correcto
		1	2	
Paso 0	Número inicial de casos 1	352	0	100,0
	2	40	0	0,0
Porcentaje global				89,8

a. En el modelo se incluye una constante.

b. El valor de corte es ,500

Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-2,175	,167	169,878	1	,000	,114

Variables que no están en la ecuación

			Puntuación	gl	Sig.
Paso 0	Variables	V005_INGRESOS	3,789	4	,435
		V005_INGRESOS(1)	1,507	1	,220
		V005_INGRESOS(2)	1,460	1	,227
		V005_INGRESOS(3)	,804	1	,370
		V005_INGRESOS(4)	,209	1	,648
	Estadísticos globales		3,789	4	,435

Bloque 1: Método =

Introducir

Pruebas omnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	3,988	4	,408
	Bloque	3,988	4	,408
	Modelo	3,988	4	,408

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	254,375 ^a	,010	,021

a. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	3	1,000

Tabla de contingencias para la prueba de Hosmer y Lemeshow

		Número inicial de casos = 1		Número inicial de casos = 2		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	34	34,000	2	2,000	36
	2	83	83,000	6	6,000	89
	3	44	44,000	4	4,000	48
	4	116	116,000	17	17,000	133
	5	75	75,000	11	11,000	86

Tabla de clasificación^a

Observado		Pronosticado		
		Número inicial de casos		Porcentaje correcto
		1	2	
Paso 1	Número inicial de casos 1	352	0	100,0
	2	40	0	0,0
	Porcentaje global			89,8

a. El valor de corte es ,500

Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	V005_INGRESOS			3,656	4	,455			
	V005_INGRESOS(1)	,206	,842	,060	1	,806	1,229	,236	6,395
	V005_INGRESOS(2)	,913	,773	1,396	1	,237	2,491	,548	11,325
	V005_INGRESOS(3)	,914	,796	1,317	1	,251	2,493	,524	11,867
	V005_INGRESOS(4)	,435	,896	,236	1	,627	1,545	,267	8,942
	Constante	-2,833	,728	15,162	1	,000	,059		

a. Variable(s) introducida(s) en el paso 1: V005_INGRESOS.

Anexo 8: Regresión Binomial Hipótesis 2

Resumen del procesamiento de los casos

Casos no ponderados ^a	N	Porcentaje
Casos incluidos en el análisis	392	100,0
Casos perdidos	0	0,0
Total	392	100,0
Casos no seleccionados	0	0,0
Total	392	100,0

a. Si está activada la ponderación, consulte la tabla de clasificación para ver el número total de casos.

Codificación de la variable dependiente

Valor original	Valor interno
1	0
2	1

Codificaciones de variables categóricas

	Frecuencia	Codificación de parámetros				
		(1)	(2)	(3)	(4)	
Los ingresos mensuales de su familia son de:	Hasta 350 dólares	89	1,000	0,000	0,000	0,000
	Hasta 600 dólares	133	0,000	1,000	0,000	0,000
	Hasta 1000 dólares	86	0,000	0,000	1,000	0,000
	Hasta 1.500 dólares	48	0,000	0,000	0,000	1,000
	Más de 1.500 dólares	36	0,000	0,000	0,000	0,000

Bloque 0: Bloque inicial

Tabla de clasificación^{a,b}

Observado			Pronosticado		Porcentaje correcto
			Número inicial de casos		
			1	2	
Paso 0	Número inicial de casos	1	307	0	100,0
		2	85	0	0,0
	Porcentaje global				78,3

a. En el modelo se incluye una constante.

b. El valor de corte es ,500

Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-1,284	,123	109,783	1	,000	,277

Variables que no están en la ecuación

		Puntuación	gl	Sig.
Paso 0 Variables	V005_INGRESOS	12,014	4	,017
	V005_INGRESOS(1)	5,895	1	,015
	V005_INGRESOS(2)	1,569	1	,210
	V005_INGRESOS(3)	2,513	1	,113
	V005_INGRESOS(4)	1,804	1	,179
	Estadísticos globales	12,014	4	,017

Bloque 1: Método = Introducir

Pruebas omnibus sobre los coeficientes del modelo

	Chi cuadrado	gl	Sig.
Paso 1 Paso	12,170	4	,016
Bloque	12,170	4	,016
Modelo	12,170	4	,016

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	397,764 ^a	,031	,047

a. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	3	1,000

Tabla de contingencias para la prueba de Hosmer y Lemeshow

		Número inicial de casos = 1		Número inicial de casos = 2		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	78	78,000	11	11,000	89
	2	109	109,000	24	24,000	133
	3	62	62,000	24	24,000	86
	4	34	34,000	14	14,000	48
	5	24	24,000	12	12,000	36

Tabla de clasificación^a

Observado			Pronosticado		Porcentaje correcto
			Número inicial de casos		
			1	2	
Paso 1	Número inicial de casos	1	307	0	100,0
		2	85	0	0,0
	Porcentaje global				78,3

a. El valor de corte es ,500

Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	V005_INGRESOS			11,535	4	,021			
	V005_INGRESOS(1)	-1,266	,478	7,004	1	,008	,282	,110	,720
	V005_INGRESOS(2)	-,820	,419	3,825	1	,050	,440	,194	1,002
	V005_INGRESOS(3)	-,256	,428	,358	1	,549	,774	,335	1,790
	V005_INGRESOS(4)	-,194	,475	,167	1	,683	,824	,324	2,090
	Constante	-,693	,354	3,844	1	,050	,500		

a. Variable(s) introducida(s) en el paso 1: V005_INGRESOS.

Anexo 9: Regresión Binomial Hipótesis 3

Resumen del procesamiento de los casos

Casos no ponderados ^a	N	Porcentaje
Casos incluidos en el análisis	392	100,0
Casos perdidos	0	0,0
Total	392	100,0
Casos no seleccionados	0	0,0
Total	392	100,0

a. Si está activada la ponderación, consulte la tabla de clasificación para ver el número total de casos.

Codificación de la variable dependiente

Valor original	Valor interno
,00	0
1,00	1

Codificaciones de variables categóricas

		Frecuencia	Codificación de parámetros (1)
Número inicial de casos	1	352	1,000
	2	40	0,000

Bloque 0: Bloque inicial

Tabla de clasificación^{a,b}

Observado	Pronosticado			
	Materias Reprobadas V56-V57		Porcentaje correcto	
	,00	1,00		
Paso 0 Materias Reprobadas V56-V57	,00	305	0	100,0
	1,00	87	0	0,0
Porcentaje global				77,8

a. En el modelo se incluye una constante.

b. El valor de corte es ,500

Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-1,254	,122	106,514	1	,000	,285

Variables que no están en la ecuación

			Puntuación	gl	Sig.
Paso 0	Variab	QCL_AP_2(1)	,726	1	,394
	Estadísticos globales		,726	1	,394

Bloque 1: Método = Introducir

Pruebas omnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	,694	1	,405
	Bloque	,694	1	,405
	Modelo	,694	1	,405

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	414,317 ^a	,002	,003

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	0,000	0	

Tabla de contingencias para la prueba de Hosmer y Lemeshow

		Materias Reprobadas V56-V57 = ,00		Materias Reprobadas V56-V57 = 1,00		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	276	276,000	76	76,000	352
	2	29	29,000	11	11,000	40

Tabla de clasificación^a

			Pronosticado		Porcentaje correcto
			Materias Reprobadas V56-V57		
Observado			,00	1,00	
Paso 1	Materias Reprobadas V56-V57	,00	305	0	100,0
	Porcentaje global	1,00	87	0	0,0
					77,8

a. El valor de corte es ,500

Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	QCL_AP_2(1)	-,320	,377	,721	1	,396	,726	,347	1,520
	Constante	-,969	,354	7,494	1	,006	,379		

a. Variable(s) introducida(s) en el paso 1: QCL_AP_2.

Anexo 10: Regresión Binomial Hipótesis 4

Resumen del procesamiento de los casos

Casos no ponderados ^a		N	Porcentaje
Casos seleccionados	Incluidos en el análisis	392	100,0
	Casos perdidos	0	0,0
	Total	392	100,0
Casos no seleccionados		0	0,0
Total		392	100,0

a. Si está activada la ponderación, consulte la tabla de clasificación para ver el número total de casos.

Codificación de la variable dependiente

Valor original	Valor interno
,00	0
1,00	1

Bloque 0: Bloque inicial

Tabla de clasificación^{a,b}

Observado		Pronosticado		
		Materias Reprobadas V56-V57		Porcentaje correcto
		,00	1,00	
Paso 0	Materias ,00	305	0	100,0
	Reprobadas 1,00	87	0	0,0
	Porcentaje global			77,8

a. En el modelo se incluye una constante.

b. El valor de corte es ,500

Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-1,254	,122	106,514	1	,000	,285

Variables que no están en la ecuación

	Puntuación	gl	Sig.
Paso 0 Variables QCL_EN_2	1,299	1	,254
Estadísticos globales	1,299	1	,254

Bloque 1: Método = Introducir

Pruebas omnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	1,352	1	,245
	Bloque	1,352	1	,245
	Modelo	1,352	1	,245

Resumen del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	413,659 ^a	,003	,005

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	0,000	0	

Tabla de contingencias para la prueba de Hosmer y Lemeshow

		Materias Reprobadas V56-V57 = ,00		Materias Reprobadas V56-V57 = 1,00		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	70	70,000	15	15,000	85
	2	235	235,000	72	72,000	307

Tabla de clasificación^a

Observado			Pronosticado		Porcentaje correcto
			Materias Reprobadas V56-V57		
			,00	1,00	
Paso 1	Materias Reprobadas V56-V57	,00	305	0	100,0
		1,00	87	0	0,0
	Porcentaje global				77,8

a. El valor de corte es ,500

Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	QCL_EN_2	-,358	,315	1,290	1	,256	,699	,377	1,296
	Constante	-,825	,392	4,437	1	,035	,438		

a. Variable(s) introducida(s) en el paso 1: QCL_EN_2.