FACULTAD DE INGENIERÍA

CARRERA PROFESIONAL DE INGENIERÍA DE SISTEMAS

TESIS

"IMPLEMENTACIÓN DE KDD PARA MEJORAR EL PROCESO DE INDENTIFICACION DE ESTILOS DE APRENDIZAJE EN LA UNIVERSIDAD AUTÓNOMA DEL PERÚ"

PARA OBTENER EL TÍTULO DE

INGENIERO DE SISTEMAS

AUTOR

FRANCISCA OFELIA RODRÍGUEZ MIRANDA

ASESOR

MG. JOSÉ LUIS HERRERA SALAZAR

LIMA, PERÚ, OCTUBRE DE 2018

DEDICATORIA

Dedico esta tesis de pre grado a mis padres, que siempre me apoyaron incondicionalmente para poder llegar hacer un profesional.

A mi esposo y mi hijo por ser el motor de culminar esta etapa en mi vida tanto en la parte moral y económica en toda mi carrera universitaria.

AGRADECIMIENTO

Primeramente, agradezco a la Universidad Autónoma del Perú, por haberme aceptado ser parte de ella, así como también a los diferentes docentes que brindaron sus conocimientos y su apoyo para seguir adelante día a día.

Agradezco también a mi Asesor de Tesis el Ing. José Luis Herrera Salazar por haberme brindado la oportunidad de recurrir a su capacidad y conocimiento científico, así como también haberme tenido la paciencia del mundo para guiarme durante todo el desarrollo de la tesis de pre grado.

Y para finalizar también agradezco a todos mis compañeros de clase durante todos los niveles de Universidad ya que gracias al compañerismo, amistad y apoyo moral han aportado en un alto porcentaje a mis ganas de seguir adelante en mi carrera profesional.

RESUMEN

La presente investigación tuvo como objetivo, descubrir cuáles son los estilos de aprendizaje dominante de estudiantes de primero y segundo ciclo de la Carrera de Ingeniería de Sistemas de la Universidad Autónoma del Peru, semestre académico 2017-1, con el propósito de mejorar el proceso de enseñanza de estudiantes.

Para llevar este estudio, se plantearon los siguientes procesos: Elaboración del estado del arte en torno a estilos de aprendizaje para conocer las diversas metodologías y resultados. El modelo utilizado, es el propuesto por Felder-Silverman, porque, este instrumento se basa en la calidad, confiabilidad y validez, para descubrir el estilo de aprendizaje dominante.
Para esta investigación se aplicó el Proceso de Extracción de Conocimiento a partir de Datos, con sus etapas: Preprocesamiento, Minería de Datos y Postprocesamiento, con el objetivo de identificar grupos homogéneos dentro de los estudiantes encuestados, por medio de técnicas descriptivas. Se utilizó software libre y de código abierto, Weka, porque permite implementar una variedad de algoritmos que después se implementó en un software que determina en tiempo real el estilo de aprendizaje dominante de los estudiantes de la Carrera.
En particular, se han propuestos algoritmos de Clustering para el análisis de la data.

Palabras clave: Estilos de aprendizaje, modelo de Felder, clúster, minería de datos, WEKA, Elearning 3.0.

ABSTRACT

The objective of this research was to discover the learning styles of the first and second cycle of the systems engineering career of the Autonomous University of Peru, academic semester 2017-1, with the purpose of improving the teaching process of the students.

To carry out this study, the following processes are proposed: Elaboration of the state of the art around learning styles to know the different methodologies and results. The model used is the one proposed by Felder-Silverman, because this instrument is based on quality, reliability and validity, to discover the dominant learning style.

For this research, the Process of Extraction of Knowledge from Data was applied, with its stages: Preprocessing, Data Mining and Postprocessing, with the objective of identifying homogeneous groups within the surveyed students, by means of descriptive techniques. It is free software and open source, Weka, because it allows implementing a variety of algorithms that are then implemented in a software that determines in real time the dominant learning style of the students of the Career.

.

In particular, grouping algorithms have been proposed for the analysis of the data.

Keywords: Learning styles, Felder model, cluster, data mining, WEKA, Elearning 3.0.

ÍNDICE DE CONTENIDO

DEDICATORIA i

AGRADECIMIENTO ii

RESUMEN iii

ABSTRACT iv

INTRODUCCIÓN xii

CAPÍTULO I. PLANTEAMIENTO METODOLÓGICO

1.1. PLANTEAMIENTO DEL PROBLEMA 2

1.1.1. Situación problemática 2

1.1.2. Descripción del problema 5

1.1.3. Enunciado del Problema 9

1.2. JUSTIFICACIÓN DE LA INVESTIGACIÓN 9

1.3. TIPO Y NIVEL DE LA INVESTIGACIÓN 11

1.3.1. Tipo de Investigación 11

1.3.2. Nivel de Investigación 11

1.4. OBJETIVOS DE LA INVESTIGACIÓN 11

1.4.1. Objetivo General 11

1.4.2. Objetivos Específicos 11

1.5. HIPÓTESIS GENERAL 12

1.6. VARIABLES E INDICADORES 12

1.7. LIMITACIONES DE LA INVESTIGACIÓN 13

1.8. DISEÑO DE LA INVESTIGACIÓN 13

1.9. TÉCNICAS E INSTRUMENTOS PARA RECOLECCIÓN DE INFORMACIÓN 14

1.9.1. Técnicas e Instrumentos 14

CAPÍTULO II. MARCO REFERENCIAL

2.1. ANTECEDENTES DE LA INVESTIGACIÓN 17

2.2. MARCO TEORICO 26

2.2.1. KDD 26

2.2.2. Modelo de Estilos de Aprendizaje 31

2.2.3. Instrumentos o Test Aplicados 32

2.2.4. CRISP-DM 40

2.2.5. Toma de Decisiones Académicas 50

CAPÍTULO III. CONSTRUCCIÓN DE LA SOLUCIÓN

3.1. ESTUDIO DE FACTIBILIDAD 56

3.1.1. Factibilidad Técnica 56

3.1.2. Factibilidad Operativa 56

3.1.3. Factibilidad Económica 57

3.2. FASE DE COMPRENSIÓN DEL NEGOCIO 58

3.2.1. Determinar objetivos del negocio 58

3.2.2. Objetivos del Negocio 59

3.2.3. Resultados esperados 59

3.2.4. Factores críticos de éxito del negocio 60

3.2.5. Inventario de recursos. 60

3.2.6. Fuentes de información 60

3.2.7. Recursos humanos 60
3.2.8. Requerimientos, expectativas y restricciones 60
3.2.9. Riesgos y contingencias 61
3.2.10. Costos y beneficios 61
3.2.11. Determinar los objetivos de DM 61
3.3. FASE DE COMPRENSIÓN Y PREPARACIÓN DE LOS DATOS 61
3.4. FASE DE MODELADO 62
3.4.1. Seleccionar técnica de modelado 62
3.4.2. Construir el modelo 62
3.5. FASE DE EVALUACIÓN 66
3.5.1. Evaluar los resultados 66
3.6. FASE DE IMPLANTACIÓN 69
3.6.1. Plan de implantación 69

CAPÍTULO IV. ANÁLISIS DE RESULTADOS Y CONTRASTACIÓN DE LA HIPÓTESIS

4.1. POBLACIÓN Y MUESTRA 73
4.1.1 Población 73
4.1.2 Muestra 73
4.2. NIVEL DE CONFIANZA Y GRADO DE SIGNIFANCIA 73
4.3. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS 73
4.3.1. Resultados Genéricos 73
4.3.2. Resultados de Clústeres 73
4.4. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS 83

4.5. CONTRASTACIÓN DE LA HIPÓTESIS 86

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES 92

5.2. RECOMENDACIONES 93

REFERENCIAS BIBLIOGRÁFICAS

ANEXOS

GLOSARIO DE TERMINOS

ÍNDICE DE TABLAS

Tabla 1 Datos actuales de los indicadores 9

Tabla 2 Descripción del indicador de la variable independiente 12

Tabla 3 Descripción de los indicadores de la variable dependiente. 12

Tabla 4 Operacionalización del indicador de la variable independiente. 13

Tabla 5 Operacionalización de los indicadores de la variable dependiente. 13

Tabla 6 Técnicas e instrumentos de la investigación de campo 14

Tabla 7 Técnicas e instrumentos de la Investigación documental 15

Tabla 8 Fases del Algoritmo en la Minería de Datos 30

Tabla 9 Dimensiones de estilos de aprendizaje de Felder 32

Tabla 10 Hoja de calificación del cuestionario de Felder (ILS) 33

Tabla 11 Perfil de un estudiante modelo de Felder 34

Tabla 12 Adaptación de Contenidos 40

Tabla 13 Factibilidad Económica 57

Tabla 14 Comportamiento de los estudiantes de acuerdo con su aprendizaje 66

Tabla 15 Correspondencia entre los estilos de aprendizaje respuestas del test de Felder 67

Tabla 16 Representación de los atributos por los estilos correspondientes 67

Tabla 17 Combinación de estilos de aprendizaje representada por cada clúster 68

Tabla 18 Cantidad de Estudiantes por Clúster de la Sección A 73

Tabla 19 Estilos de Aprendizaje del Clúster 1 73

Tabla 20 Estilos de aprendizaje del Clúster 2 74

Tabla 21 Cantidad de Estudiantes por Clúster de la Sección B 74

Tabla 22 Estilos de aprendizaje del Clúster 1.....74
Tabla 23 Estilos de aprendizaje del Clúster 2.....75
Tabla 24 Cantidad de Estudiantes por clúster de la Sección C 75
Tabla 25 Estilos de aprendizaje del Clúster 1.....75
Tabla 26 Estilos de aprendizaje del Clúster 2.....76
Tabla 27 Cantidad de estudiantes por clúster de la Sección D.....76
Tabla 28 Estilos de aprendizaje del Clúster 1.....76
Tabla 29 Estilos de aprendizaje del Clúster 2.....77
Tabla 30 Cantidad de estudiantes por clúster de la Sección F 77
Tabla 31 Estilos de aprendizaje del Clúster 1.....77
Tabla 32 Estilos de aprendizaje del Clúster 2.....78
Tabla 33 Resultados de evaluación parciales 78
Tabla 34 Porcentaje de aprobados y desaprobados por sección 79
Tabla 35 Porcentaje de alumnos aprobados y desaprobados por curso 80
Tabla 36 Datos de la Preprueba 81
Tabla 37 Datos de la postprueba 82
Tabla 38 Promedio de los indicadores de la preprueba y Post Prueba 83
Tabla 39 Resultados de Pre-Prueba y Post Prueba para el KPI1.....83
Tabla 40 Resultados de Pre-Prueba y Post-Prueba para el KPI2 84
Tabla 41 Contratación de la Hipótesis.....86
Tabla 42 Resumen de prueba t de student del KPI1 88
Tabla 43 Resumen de prueba t de student del KPI1 90

ÍNDICE DE FIGURAS

Figura 1 Ubicación de la Universidad Autónoma del Perú 4

Figura 2 Proceso Actual. 7

Figura 3 Proceso de Toma de Decisiones. 8

Figura 4 Proceso del KDD. 26

Figura 5 Resultados de MEMLETICS 38

Figura 6 Ciclo de Vida de CRISP-DM 40

Figura 7 Fase de Comprensión del Negocio 42

Figura 8 Fase de Comprensión de los Datos. 43

Figura 9 Fase de Preparación de Datos 45

Figura 10 Fase de Modelado. 46

Figura 11 Fase de Evaluación 48

Figura 12 Fase de Implementación 50

Figura 13 Proceso de Toma de Decisiones 51

Figura 14 Combinación de estilos de aprendizaje por clúster 68

Figura 15 Esquematización preliminar de las tres primeras fases del proceso de adaptación 69

Figura 16 Página Web para generar estilos de aprendizaje por sección 70

Figura 17 Página Web para mostrar los estilos de aprendizaje por sección 71

Figura 18 Diagrama de alumnos aprobados y desaprobados por sección 79

Figura 19 Distribución de probabilidades 87

Figura 20 Distribución de Probabilidad del KPI1 89

INTRODUCCIÓN

En el ambiente estudiantil, uno de los procesos que a los estudiantes les permita evidenciar su aprendizaje en clases o fuera de ella es por medio de la realización de sus trabajos o tareas. Sin embargo, en este proceso se debe considerar que no todos los estudiantes logran su aprendizaje de una manera satisfactoria, es decir, no todos los estudiantes tienen un mismo estilo de aprendizaje o terminan aprendiendo todos en un mismo nivel, lo que esta sucediendo en la carrera de ingeniería de Sistemas.

Si los estudiantes tienen sus formas de aprendizaje, el docente también tiene su forma de enseñanza, como: leer, debatir, reflexionar, mientras ciertos docentes se centran en la práctica más que la teoría o el análisis de casos. Pero hay un problema, aún no hemos podido comprender del todo cómo nuestros estudiantes aprenden, tan solo conocemos que existen muchas maneras de aprender, tales como: observando, escuchando, reflexionando, debatiendo y hasta a veces actuando. Pero existen otras, que requieren de lógica y razonamiento para aprender, como: memorizando, visualizando, construyendo analogías y hasta a veces de forma intuitiva.

Es así que el objetivo de esta tesis es la Identificar el estilo de aprendizaje tratado de una forma técnica, para ello se plantea la aplicación de técnicas de minería de datos, mediante la técnica de aprendizaje no supervisado, es decir no existe un conocimiento a priori, donde se analicen las relaciones entre variables.

Para esto la investigación se ha divido en varios capítulos que detallo a continuación:

En el capítulo I. Planteamiento metodológico. - se detalla la verdad circunstancia y las delimitaciones de la exploración, para después seguir con la definición del inconveniente, los objetivos y conjetura. Además, se define el tipo y nivel de exploración, de esta forma como el procedimiento y diseño de la misma.

En el capítulo II. Marco referencial se describen los antecedentes de la exploración, después se hace una regencia histórica del tema a tratar y al final se lleva a cabo el marco conceptual.

En el capítulo III. Construcción del sistema KDD se describe la implementación de la herramienta a nivel prototipo.

En el capítulo IV. Análisis e interpretación de resultados se ejecuta la prueba empírica para la recolección, análisis e interpretación de los resultados que se consiguieron, en el contexto de la prueba de la conjetura. Ahora, se procede a hacer la articulo prueba para los grupos seleccionados, en rigurosa aplicación de interfaz elegido, para corroborar la validez de la conjetura de exploración.

En el Capítulo V. Las conclusiones y recomendaciones pertinentes de los resultados conseguidos, además se formulan las sugerencias, visualizaciones, críticas y comentarios sobre los resultados de los indicadores sobre la utilización de la utilidad y la circunstancia de hoy.

Las referencias bibliográficas se listan los artículos de revistas, libros y otras monografías, informe científico o técnico, tesis doctoral y monografías en internet.

Al final, la implementación de la red social usando metodologías ágiles para el proceso de participación estudiantil en la Universidad Autónoma del Perú; esto fue posible con el apoyo y participación de todas las autoridades y personal de la institución, quienes influirán directamente en la eficiente y efectiva implementación del sistema; por lo cual se recomienda que todas las universidades implementen este sistema.

CAPÍTULO I

PLANTEAMIENTO METODOLÓGICO

1.1. PLANTEAMIENTO DEL PROBLEMA

1.1.1. Situación problemática

Realidad Mundial

Las Universidades extranjeras cuentas con grandes volúmenes de información almacenada en su base de datos la cual constituyen un activo importante en su organización, ya que estas generan procesos que soportan sus operaciones para una mejora en el servicio académico de sus estudiantes.

Esto significa que, en sus bases de datos, se encontraría las características de los estudiantes, el comportamiento académico, las estructuras productivas, los resultados de diversas estrategias y muchas claves para la mejora del desempeño. Por lo mismo cuenta con sistema de información de grandes costos, pero algunas de ellas no están mostrando los indicadores o patrones que desean y por lo general su toma de decisión no es muy asertiva.

El dominio de estas nuevas tecnologías y la gestión de información de los Procesos están basados en la explotación de datos y es un gran desafío para los colaboradores. Por lo cual parte de sus inversiones de la empresa se pierde en capacitaciones y desarrolladores para que su sistema sea más dinámico y rápida en mostrar procesos académicos por los mismo que desean que su tipo de enseñanza sea la mejor a nivel mundial

Realidad Nacional

La gran mayoría de las empresas dedicadas a este rubro no cuentan con un sistema de servicio académico dinámico y de fácil uso con un plan de control que haga que la información almacenada sea mostrada de forma rápida, lo que generaría pérdida de tiempo de los colaboradores en la

búsqueda de los registros académicos. Por otra parte, se produciría gastos en la capacitación del personal para una mejor producción.

En la provincia de Lima la estructura del sistema educativo es más desarrollada en universidades privadas que las universidades públicas, ya que los aportes que reciben ellos son mínimos y los proyectos que se realizan en ese mismo sector es de largo plazo, las pocas comisiones que reciben las universidades públicas son para su infraestructura y no para el tipo de estrategias que puedan contar para mejorar el rendimiento académico.

Asimismo, los colegios que cuentan con más experiencia en el mercado manejan estos procesos para un mejor desarrollo de la educación, por lo cual genera un ámbito más competitivo en la enseñanza.

Realidad Empresarial

El proceso académico de la Universidad Autónoma del Perú para observar el Rendimiento Académico del Estudiante no es muy detallada, el sistema no descarga una información completa la cual no serviría para una toma de decisión adecuada para cada tipo de estudiante, las entidades involucradas en el sistema son apellido y nombre del estudiante, edad, sexo, la escuela de procedencia, nombre y lugar de la empresa donde labora sus padres, desempeño académico, actividades externas, nombre de la carrera, ciclo, con el propósito que esta información al ser requerida en la base de datos por características similares o detalladas se muestre en el sistema.

Mucha de la información almacenada no pueda ser visibles en el sistema al momento de haber sido requerida por un patrón o dato específico ya que no existe un proceso de agrupación con características similares para observar los indicadores de cada estudiante.

Al no poder trabajar con este sistema que no presenta una información completa y que no se muestre los datos detallados al momento de ser agrupado por patrones crea una gran dificultad en la administración ya que al no descargar los datos pedidos la persona encargada de mostrar los indicadores diarios quincenales o mensuales deberá de buscar uno por uno en la base de datos y así agruparlos en similitudes lo cual es una pérdida de tiempo al querer tomar decisiones eficientes y al querer que las nuevas estrategias sean rápidas las cuales logren ayudar al estudiante a mejorar su rendimiento académico.

- La investigación, se realizará en la Universidad Autónoma del Perú, ubicada en la Carretera de la Panamericana Sur Km. 16.3 Mz. A Lote. 06 (Altura Del Parque Zonal Huayna Cápac – By Pass) Villa El Salvador Lima – Perú, no obstante, la aplicabilidad y el alcance de sus resultados, tienen validez en cualquier organización de su rubro.

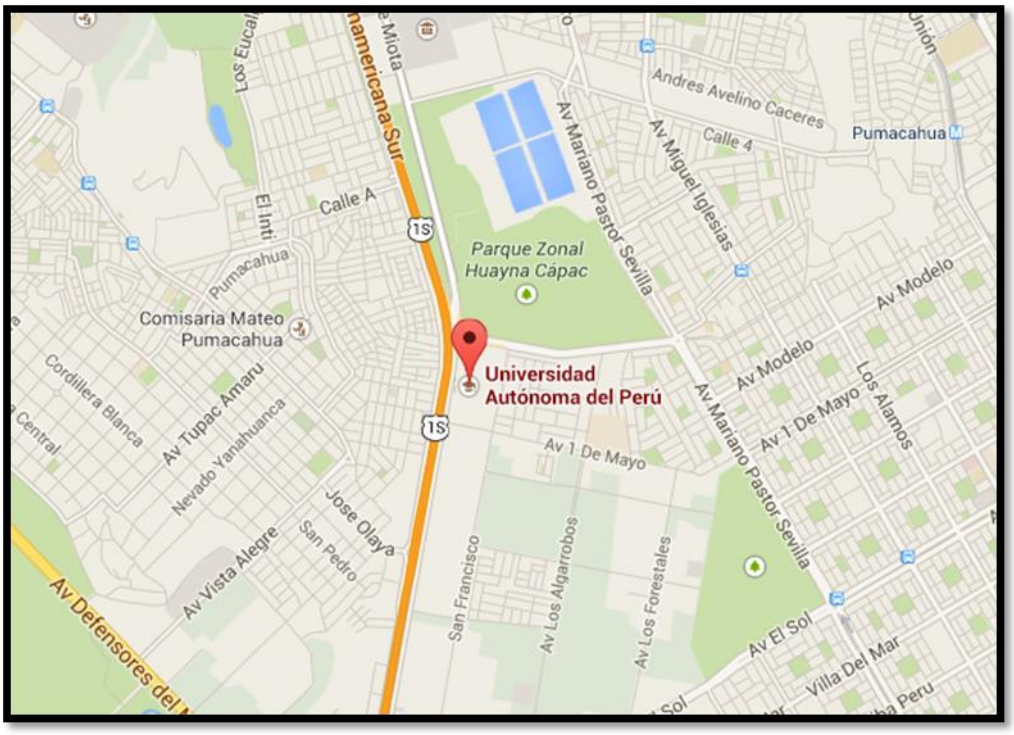


Figura 1. Ubicación de la Universidad Autónoma del Perú
Adaptado de “Ubicación de la Universidad Autónoma del Perú” por Google Maps, 2018

1.1.2. Descripción del problema

La Universidad Autónoma del Perú es una empresa nueva que está interesada en la educación y la exigencia competitiva académica de los estudiantes.

En los últimos años, los datos grabados de la empresa en la base de datos han ido incrementándose considerablemente. Esta información es de gran importancia estratégica para la Universidad, se accede a través del uso de técnicas clásicas como son sentencias SQL y los procedimientos almacenados. Por tal motivo existe demora en la recuperación y el análisis de la información para la elaboración de informes, formularios y reportes de gestión.

Así es necesaria la utilización de métodos analíticos más avanzados, como es la minería de datos para la explotación de datos con la finalidad de ayudar a los docentes en los sus estilos de enseñanza.

Por lo general hay muchos problemas que guardan interacción de confianza con el criterio de “estilos de aprendizaje”, al que los psicólogos de la educación atribuyen las formas particulares de actuar de cada individuo en el desarrollo de estudio.
Un aspecto sustancial es, que los estilos de estudio pertenecen a los pilares esenciales para desarrollar el estudiar a estudiar. Entre las cuestiones que frecuentemente los docentes nos hacemos están las siguientes:
1. ¿Cómo aprenden mis alumnos?
2. ¿Por qué aprenden cuando aprenden?
3. ¿Por qué a veces no consiguen aprender al menos en el grado en que me había propuesto?
4. ¿Por qué en el caso de dos estudiantes de la misma edad, del mismo ambiente sociocultural y con similares capacidades intelectuales, ante una misma situación de aprendizaje y dentro de un mismo contexto, uno aprende y el otro no?

5. ¿Por qué un mismo método de lectura, utilizado por un mismo profesor, puede ser causa de fracaso, frustración e incluso rechazo para algunos alumnos, mientras para otros puede resultar un método excelente?

6. ¿Por qué el rendimiento de unos aumenta cuando trabajan en equipo mientras otros necesitan del silencio y del trabajo individual para rendir al máximo?

¿Qué ventajas ofrece conocer y potenciar los estilos de aprendizaje de los estudiantes?

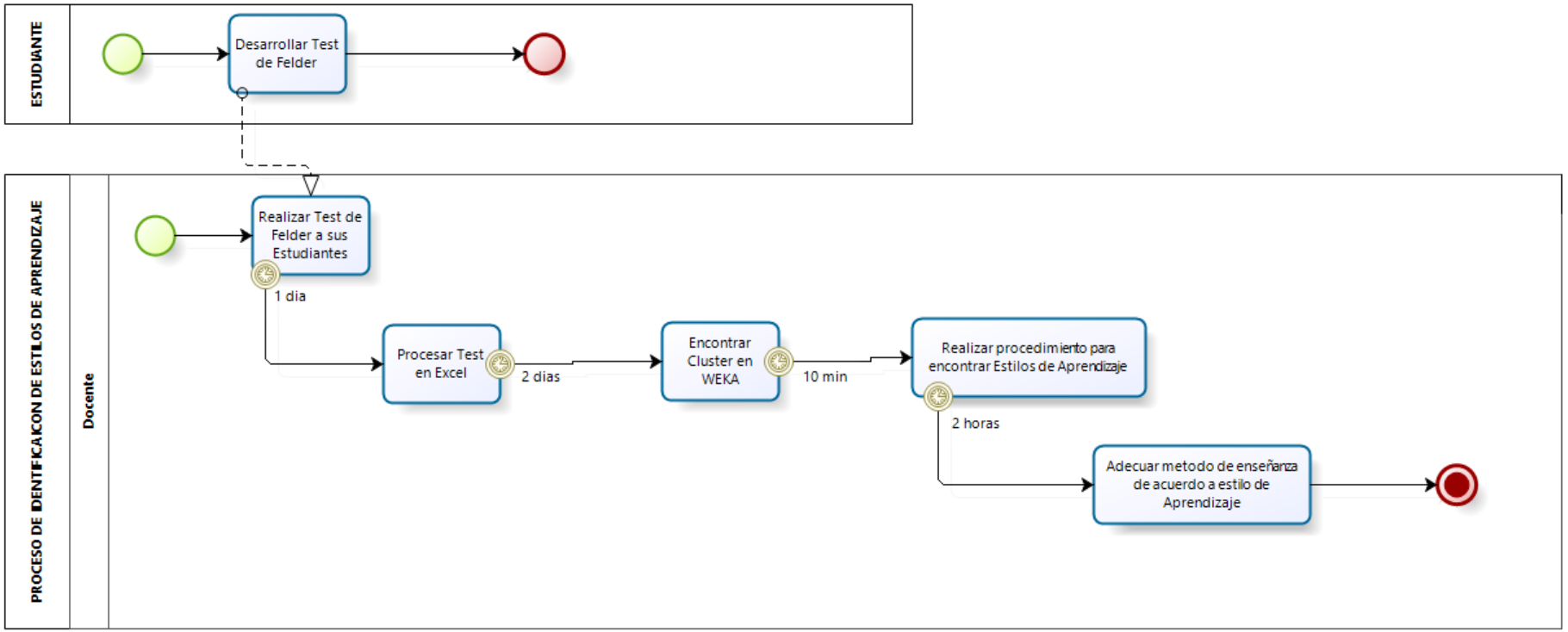


Figura 2. Proceso Actual.

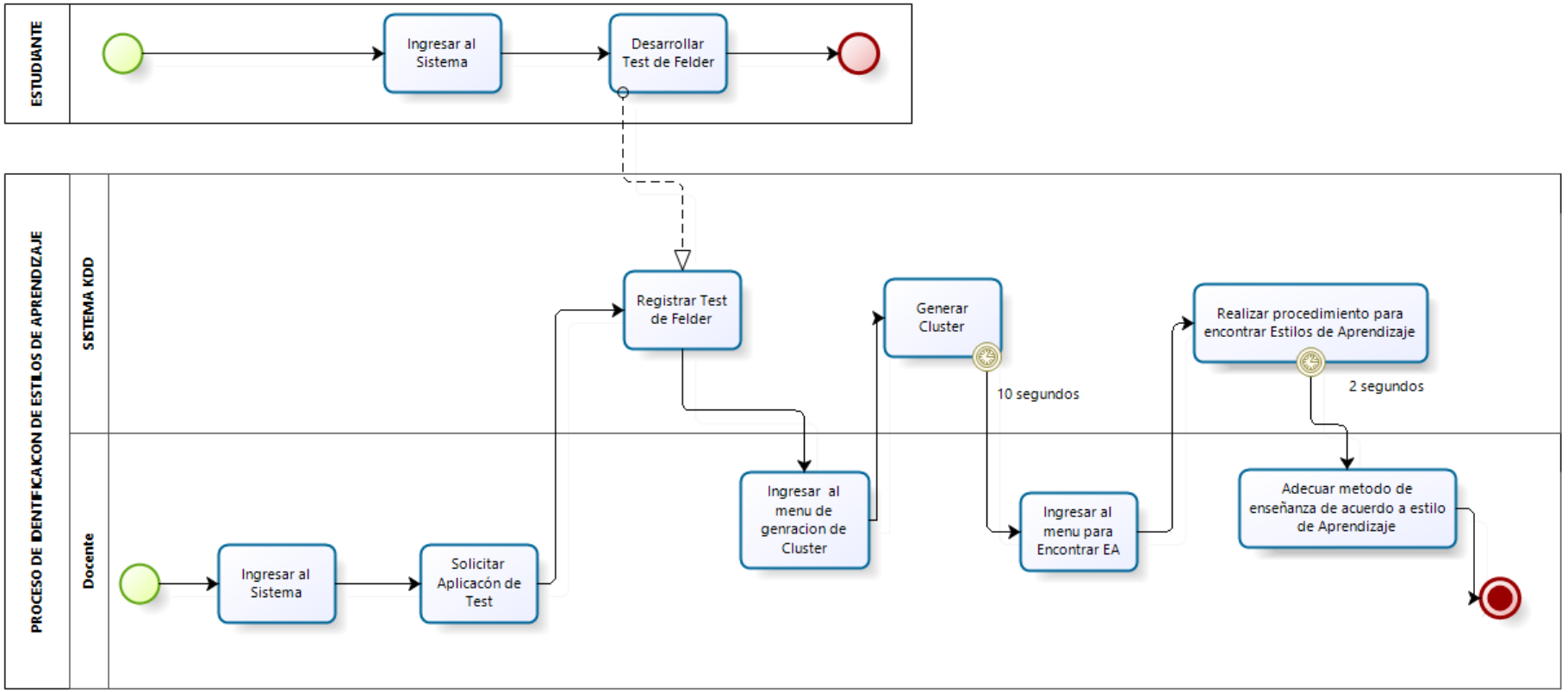


Figura 3. Proceso de Toma de Decisiones.

El siguiente proceso muestra problemas en:

- Tiempo de demora en tomar la decisión.
- Tiempo de demora en analizar los reportes.
- Cantidad de Estudiantes por estilos de aprendizaje
- Porcentaje de desertores en el año.
- Exactitud de la Información

Tabla 1
Datos actuales de los indicadores

Indicador	Datos de Pre-Prueba (Promedio)
Tiempo empleado para determinar los clústeres de estilos de aprendizaje	2 días
Cantidad de Estudiantes por estilos de aprendizaje x sección	15

1.1.3. Enunciado del Problema

¿En qué medida la implementación de minería de datos mejorará el Proceso de Identificación de Estilos de Aprendizaje en la Universidad Autónoma del Perú?

1.2. JUSTIFICACIÓN DE LA INVESTIGACIÓN

La presente Investigación se justifica en lo siguiente:

La toma de decisiones es muy importante y de ella se pueden obtener muchos beneficios, pero es necesario contar con la información correcta y adecuada, así se podrán formular estrategias que beneficien a los estudiantes.

De esta manera se podrán realizar acciones con respecto a la información proporcionada generando valor a la educación y evitando que los estudiantes abandonen la universidad.

Se tienen la posibilidad de destacar las siguientes ventajas:

a) Poder orientar mejor el estudio de cada alumno si conocemos cómo aprenden. Ósea, que la selección de nuestras tácticas didácticas y estilo de lección va a ser más eficaz.
b) La aplicación en el aula de los estilos de estudio es el sendero más científico de que se dispone para individualizar la instrucción.
c) Si la misión didáctica es conseguir que el alumno aprenda a estudiar, entonces debemos apostar por asistirlo a comprender y hacer mejor sus propios estilos de estudio.

Esto le permitiría al alumno, entre otras cosas, saber:

- ✓ Cómo controlar su propio aprendizaje.
- ✓ Cómo diagnosticar sus puntos fuertes y débiles como alumno.
- ✓ Cómo describir su estilo o estilos de aprendizaje.
- ✓ Conocer en qué condiciones aprende mejor.
- ✓ Cómo aprender de la experiencia de cada día.
- ✓ Cómo superar las dificultades que se le presentan en su proceso de aprendizaje.

Resumiendo, el saber de los estilos de estudio predominantes en las salas universitarias tiene la posibilidad de ser una utilidad docente de gran ayuda para adaptar el estilo de lección del instructor para conseguir un mejor desempeño estudiantil del alumno. A la vez permitiría crear procedimientos de evaluación más apropiados para corroborar el avance de los alumnos generalmente. De igual modo, para los alumnos sería de una enorme herramienta porque podrían planificar el estudio según sus propios estilos, evadiendo de esta forma bloqueos y mejorando sus resultados.

Relevancia Social. - La opinión de los estudiantes en relación a la educación que reciben, es muy importante, ya que esto generará un mayor conocimiento acerca del tipo de educación que brinda la universidad, por ello es necesario tener la información adecuada para poder tomar una decisión que beneficie tanto a la universidad como a los estudiantes de esta.

Entorno Empresarial. - Es necesario que la universidad conozca el tipo de estudiantes que posee, esto puede ser beneficioso no solo para la universidad sino también para el estudiante ya que este se verá beneficiado con las decisiones que tome la universidad con respecto a la educación que brinda.

Implicaciones Prácticas. - La aplicación de Minería de Datos mejorará el Proceso Académico de la Universidad Autónoma del Perú de tal manera que se evitará la disminución de estudiantes, mediante la toma de decisiones, dependiendo de las necesidades del estudiante y ayudando a que siga estudiando en la universidad

1.3. TIPO Y NIVEL DE LA INVESTIGACIÓN

1.3.1. Tipo de Investigación

Aplicada, Se aplicarán los conocimientos para poder dar solución al problema presentado sobre la Toma de Decisiones sobre temas relacionados al rendimiento académico.

1.3.2. Nivel de Investigación

Explicativa: Debido a que la investigación desarrollará las causas del proceso de toma de decisiones, así como los efectos que aportará a nuestra solución de minería de datos.

1.4. OBJETIVOS DE LA INVESTIGACIÓN

1.4.1. Objetivo General

Determinar en qué medida la implementación de minería de datos utilizando la Metodología CRISP-DM mejora el Proceso de identificación de estilos de aprendizaje de los estudiantes de la Universidad Autónoma del Perú.

1.4.2. Objetivos Específicos

- Determinar en qué medida se reduce Tiempo empleado para determinar los clústeres de estilos de aprendizaje.
- Determinar en qué medida se reduce el tiempo empleado para determinar los estilos de aprendizaje por clúster

1.5. HIPÓTESIS GENERAL

La aplicación de minería de datos utilizando la metodología CRISP-DM permite mejorar el proceso de identificación de estilos de aprendizaje en la Universidad Autónoma del Perú.

1.6. VARIABLES E INDICADORES

A. Variables

a) **Variable Independiente:** KDD

b) **Variable Dependiente:** Proceso de identificación de estilos de aprendizaje

c) **Variables Interviniente:** CRISP-DM

B. Indicadores

B.1. Conceptualización

a) **Variable Independiente:** KDD

Tabla 2

Descripción del indicador de la variable independiente

Indicador: Presencia – Ausencia
Descripción: Cuando indique NO, es porque no ha sido implementado KDD en el proceso de enseñanza- aprendizaje y aún se encuentra en la situación actual del problema. Cuando indique SI, es cuando se ha implementado KDD, esperando obtener mejores resultados.

b) Variable Dependiente: Proceso de identificación de estilos de aprendizaje

Tabla 3

Descripción de los indicadores de la variable dependiente.

Indicador	**Descripción**
Tiempo empleado para determinar los clústeres de estilos de aprendizaje	Es el tiempo para poder procesar y determinar los clústeres de estilos de aprendizaje
Tiempo empleado para determinar los estilos de aprendizaje por clúster	Es el tiempo para poder procesar y los estilos de aprendizaje de acuerdo a los clústeres predominantes

B.2. Operacionalización

a) Variable Independiente: KDD

Tabla 4
Operacionalización del indicador de la variable independiente.

Indicador	Índice
Presencia – Ausencia	No, Si

b) Variable Dependiente: Proceso de identificación de estilo de aprendizaje

Tabla 5
Operacionalización de los indicadores de la variable dependiente.

Indicador	Índice	Unidad de Medida	Unidad de Observación
Tiempo empleado para determinar los clústeres de estilos de aprendizaje	[1 - 10]	Días	Reloj
Tiempo empleado para determinar los estilos de aprendizaje por cluster	[1 - 10]	Días	Reloj

1.7. LIMITACIONES DE LA INVESTIGACIÓN

- ✓ La principal limitación es que no se brinde el permiso para poder encuestar a los estudiantes
- ✓ No contar con mucho tiempo.
- ✓ Perdida de alguna información valiosa más minuciosa.
- ✓ No contar con la información necesaria al tiempo requerido.

1.8. DISEÑO DE LA INVESTIGACIÓN

Ge O_1 **X** O_2

Donde:

Ge = Grupo experimental, conformado por el número representativo al proceso de Toma de Decisiones.

$O_1 =$ Son los valores de los indicadores de la variable dependiente en la Post-Prueba.

X = KDD

$O_2 =$ Son los valores de los indicadores de la variable dependiente en la Pre-Prueba.

Descripción:

Se trata de la conformación de un grupo experimental (Ge) conformado por el número representativo de actividades del proceso de Toma de Decisiones, al cual a sus indicadores de Pre-Prueba (O_1), se le administra un estímulo o tratamiento experimental, la implementación de KDD como estímulo (X) para solucionar el problema de dicho proceso, luego se espera que se obtenga (O_2).

1.9. TÉCNICAS E INSTRUMENTOS PARA RECOLECCIÓN DE INFORMACIÓN

1.9.1. Técnicas e Instrumentos

A. Técnicas e instrumentos de la Investigación de Campo

Tabla 6

Técnicas e instrumentos de la investigación de campo

TÉCNICAS	INSTRUMENTOS
1. Observación Directa Encargado de la Carrera Estructurada No Participante	• Reportes de Rendimiento • Sistemas Transaccionales
2. Aplicación de Cuestionarios Abierto Cerrado	• Cuestionario Test de Felder

B. Técnicas e instrumentos de la Investigación Documental

Tabla 7
Técnicas e instrumentos de la Investigación documental

TÉCNICAS	INSTRUMENTOS
Revisión de: • **Libros** • **Solicitudes de Reporte** • **Tesis** • **Artículos Científicos** • **Documentación Estadística** • **Informes de Entrega** • **Revistas** • **Internet**	• Computadora • USB • Fotocopias • Impresiones • Libreta de Apuntes • Diapositivas

CAPÍTULO II

MARCO REFERENCIAL

2.1. ANTECEDENTES DE LA INVESTIGACIÓN

TESIS

a) Autor: Juan Carlos Gonzáles Cardona

Título: SISTEMA DE APOYO PARA LA ACREDITACIÓN DE LA CALIDAD DE PROGRAMAS ACADÉMICOS DE LA UNIVERSIDAD DE CALDAS, APLICANDO TÉCNICAS EN MINERÍA DE DATOS, 2013.

Correlación:

Una de las primordiales adversidades que enfrenta el sistema educativo de hoy es la deserción, su valor juntado llega a escenarios del 45% en todo el país y del 33% en la Facultad de Caldas; una de sus primordiales causas es la deserción de tipo estudiantil, por lo cual se hace esencial la definición de un indicio que pueda medir el desempeño estudiantil en sus diferentes dimensiones (la excelencia, la eficacia y la eficacia). Una vez definido el indicio, es primordial saber los componentes que influyen en este con el fin de tomar acciones que tengan el más grande encontronazo posible; estos componentes tienen dentro elementos de identidad, socioeconómicos, vocacionales, de estudios previos, del ámbito familiar etc y sus relaciones, por lo cual un análisis multivariado debe ser el tipo de modelo que los caracterice, la regresión logística y su extenso uso en la determinación de componentes de peligro o de custodia fue el análisis elegido, tanto por esta propiedad como por su manejo de cambiantes de tipo numérico como categóricas.

Este desarrollo de extracción de conocimiento desde los datos KDD, que estuvo en apogeo en los años anteriores en los ambientes académicos, se desarrolló usando la metodología CRISP-DM y fue establecido sobre la aplicación de libre uso RapidMiner y comprende desde la extracción de información de la banco de información del Sistema de Información Académica SIA, su transformación, validación, el cálculo de los índices y el indicio de desempeño estudiantil IAR, el análisis de Regresión Logística por programa y nivel de adelanto en créditos del programa, hasta la generación de reportes de tipo descriptivo como del modelo. Los resultados finales detallan los causantes de compromiso y de custodia en el desempeño estudiantil para los alumnos de cada programa presencial en diferentes instantes de su paso por la facultad. (Gonzáles, 2013)

b)Autor: Erwin Sergio Fischer Angulo

Título: MODELO PARA LA AUTOMATIZACIÓN DEL PROCESO DE DETERMINACIÓN DE RIESGO DE DESERCIÓN EN ESTUDIANTES UNIVERSITARIOS, 2012.

Correlación:

Más allá de los esfuerzos en reglas públicas para ofrecer ingreso a la facultad, la deserción universitaria se convirtió en un inconveniente prioritario a ser investigado y tratado. La tasa de deserción llegó a constituir uno de los más importantes indicadores de eficacia interna dentro de alguna institución de educación superior. Invertir más tiempo en diagnósticos de las razones de la deserción con formas correctas que permitan adivinar ésta con más grande eficacia, ayuda a hacer mejor la relación efectividad-costo en la administración de la unidad académica.

El propósito del presente emprendimiento radica en investigar y ofrecer una metodología que permita detectar en forma automática a los alumnos con más grande compromiso de deserción de las carreras de Ingeniería de la Facultad de Las Américas. Para la utilización de este emprendimiento se adoptó la metodología CRISP-DM que composición el desarrollo de minería de datos en seis fases, que interactúan entre ellas de manera iterativa. Se aplicaron los modelos de Redes Neuronales, Árboles de elección y Cluster K-medianas para investigar el accionar de los alumnos, evaluando componentes como el puntaje promedio obtenido en la Prueba de Selección Universitaria (PSU), el promedio de notas obtenido en la lección media, La edad a la fecha de Ingreso a la institución y el género de los alumnos.

La precisión de los modelos es calculada desde el grupo de datos de pruebas, los cuales señalan que ningún modelo predictivo lanzó resultados positivos, gracias a esto se analizó el desarrollo y se llegó a la conclusión que es muy posible que los datos de entrada no eran bastante confiables. Ya que dentro de los parámetros de este trabajo era irrealizable hallar datos legítimos y terminados, esta proposición ofrece una metodología para combatir estudios de minería de datos didáctica donde se eviten los inconvenientes descritos Como trabajo futuro se ofrece llevar a cabo un desarrollo de captura de cambiantes importantes para la exploración, al instante del ingreso del estudiante a la

institución superior, de tal forma de hacer más simple la generación de un comercio de datos para contribuir a la toma de elecciones. (Fischer, 2012)

c) Autor: Jesús Germán Andrés Pautsch

Título: MINERÍA DE DATOS APLICADA AL ANÁLISIS DE LA DESERCIÓN EN LA CARRERA DE ANALISTA EN SISTEMAS DE COMPUTACIÓN, 2009.

Correlación:

En los años anteriores, la tasa mundial promedio de desarrollo del área informático fue del 20% anual, en tanto que en la Argentina fue menor al 8 % anual. Este porcentaje tiene la posibilidad de ser más grande si egresaran de las universidades más expertos informáticos, más que nada de las carreras a corto período que son las que las compañías más denuncian.

En esta proposición de nivel se llevó a cabo una Minería de Datos sobre el Cubo 04 Desgranamiento, exportado del Sistema de Administración Académica SIU-Guaraní, provistos por el Ministerio de Educación, Ciencia y Tecnología de la Nación. Se ha experimentado con procedimientos Supervisados, como la Categorización y No Supervisados como es el Agrupamiento.

El propósito primordial fue maximizar la calidad que los modelos tienen para clasificar y agrupar a los alumnos, según sus propiedades académicas, causantes sociales y demográficos, que han desertado de la Carrera Analista en Sistemas de Computación de la Facultad de Ciencias Exactas, Químicas y Naturales de la Facultad Nacional de Misiones.

El trabajo se desarrolló bajo la metodología de libre difusión Crisp-DM y con la utilidad comercial IBM DB2 Warehouse (versión 9.5). Los resultados que se consiguieron permitieron ver, no sólo los distintos patrones de los estudiantes desertores, sino que además las propiedades de los alumnos en actividad y egresados.

La calidad de los modelos obtenidos por medio de la ordenación con árboles de elección superó a la técnica de agrupamiento por medio de la generación de clústeres y las dos han superado extensamente lo planteado. Más allá de que los

profesionales en el estudio de la deserción universitaria han elogiado la aptitud que los modelos tienen para distinguir a cada tipo de alumno, han planteado la necesidad de contemplar más cambiantes colegas económicas en el estudio de la deserción. (Andrés, 2009)

d) Autor: Olga Conde Gómez

Título: "RELACIÓN ENTRE ESTILOS DE APRENDIZAJE AUTOESTIMA Y RENDIMIENTO ESCOLAR EN ESTUDIANTES DEL PRIMER AÑO DE BACHILLERATO", 2015.

Correlación:

El trabajo llega a la conclusión de que el estilo de aprendizaje asimilador de Kolb predomina en los estudiantes de Primer año de bachillerato tanto en la variante científico tecnológica y científico humanista, con respecto al rendimiento escolar, este es superior en las mujeres que, en los hombres, pues ellas destacan en los niveles bueno y muy bueno de la variable. (Conde, 2005)

LIBROS

a) Autor: Ing. Guillermo E. Gil Albarran

Título: DATA MINING - MINERÍA DE DATOS Y SQL, 2009.

Correlación:

Este libro muestra la minería de datos con un enfoque eminentemente efectivo puesto que el contenido de cada capítulo empieza exponiendo los conceptos correctos, ilustrándolos con ejemplos servibles para ahora detallar las diferentes técnicas de minería de datos y temas, así como la selección de datos y muestras, la navegación y análisis amigable de datos y su depuración y modificación.

El KDD aparece como una tecnología que intenta contribuir a abarcar el contenido de una base de datos de información. De manera general, los datos son la materia prima bruta. En el instante que el usuario les asigna algún concepto particular pasan a transformarse en información. Mientras avanza el contenido del libro el cual son VIII capítulos se tratan técnicas particulares de KDD, como son el clustering, los modelos, los árboles de elección, las redes neuronales, las técnicas de predicción, de reducción de la extensión y de ordenación. Nuestra aptitud para guardar datos creció en los años anteriores a

velocidades exponenciales. En contrapartida, nuestra aptitud para procesarlos y utilizarlos no fue a la par. Por esto, el KDD, se muestra como una tecnología de acompañamiento para examinar, investigar, abarcar y utilizar el saber obtenido utilizando enormes volúmenes de datos. Conocer nuevos caminos que nos ayuden en la identificación de atrayentes construcciones en los datos se encuentra dentro de las tareas esenciales en el KDD.

En el tema comercial, resulta atrayente hallar patrones ocultos de consumo de los usuarios para lograr examinar nuevos horizontes. Este Libro nos transporta a niveles donde tenemos la posibilidad de utilizar construcciones de minería de datos como la de Bike Buyer la cual es un banco de información estratégica que exhibe cómo hacer, entrenar y examinar modelos de minería de datos del árbol de elección y clústeres usando el lenguaje de consulta Extensiones de minería de datos llamada (DMX). (Gil, 2009)

b) Autor: Luis Paulo Vieira, Luis Iván Ortiz valencia, Santiago Segundo Ramírez

Título: INTRODUCCIÓN A LA MINERÍA DE DATOS, 2009.

Correlación:

Antes de llevar a cabo un sistema con KDD debemos entender de la consideración y conceptos básicos de lo que es Minería de Datos y este libro nos asiste a comprender lo servible y bueno puede ser tener una masa de datos y intentar llevar las con técnicas con el propósito de hacer mejor el desarrollo de información.

Realizando un corto resumen la minería de datos se enseña cómo desarrollo de extracción de conocimiento desde datos. Este desarrollo se compone de numerosas fases y tiene dentro diferentes técnicas de los campos del estudio automático. Por medio de varios ejemplos completados en packs de minería de datos, las técnicas que se necesitan en cada etapa del desarrollo son: técnicas de elaboración y almacenes de datos, técnicas propias de extracción de modelos (clasificación, agrupamiento, regresión, organización, etc. y técnicas de evaluación y difusión del conocimiento extraído.

El propósito del libro de notificarnos con la Minería de datos es lograr y conocer un modelo que sea con la capacidad de hallar información servible escondida al interior de las bases de datos. El lenguaje de conjetura detalla el modelo. Como todas las conjeturas consisten en las mismas tablas de diccionarios de datos, atributos y relaciones. Dicen que no existe sólo una forma o técnica de llevar a cabo estudio de máquinas o reconocimiento de patrones, diferentes teóricos asumen diferentes formas de combatir el inconveniente (sin que una sea siempre mejor que otra). Hay oportunidades en que se requiere utilizar diferentes técnicas en numerosas etapas del análisis del procesamiento de datos. Esto ilustra la validez del ámbito de “estrategias múltiples” que ofrece la minería de datos.

Unas de las varias herramientas de minería de datos como SPSS Clementine, WEKA y otras son ilustradas en el libro, las implementaciones de cada una de ella están aplicadas en un sistema y están se explican de manera apropiada. (Vieira, Ortiz y Ramírez, 2009)

ARTÍCULOS

a) **Autor: Alex Fernando Bonilla Gordillo, Miguel Ángel Ojeda Schuldt, Fabricio Echevarría Briones**

Título: IMPLEMENTACIÓN DE MINERÍA DE DATOS BASADA EN REDES BAYESIANAS PARA LA TOMA DE DECISIONES EN LOS REGISTROS ACADÉMICOS, 2006.

Correlación:

En el presente reporte no enseña cómo sería el sistema de un colegio aplicando la minería de datos fundamentadas en redes bayesianas la cual provee una manera compacta de representar el saber y procedimientos flexibles de razonamiento basados en teorías probabilísticas la cual es llamada Teorema de Bayes que tienen la capacidad de adivinar el valor de las cambiantes no observadas y argumentar las observadas.

Antes de emprender con el tema se detalló en el trabajo distintas formas que no cumplían algunos requerimientos como la situacion de la metodología OLAP la cual no generaba patrones, reglas ni pautas lo cual era primordial para hacer un diagnostico de forma sencilla los registros académicos y de esto poder tomar elecciones. Las resoluciones que muestran son posibilidades de eventos que se

relacionan a las ocasiones deseadas; la cual en la pregunta debe existir la interrogante y la variable a cuestionar, entre otras cosas ¿Cuál sería la posibilidad de que alguien se registre en la materia X? al agregarse un más grande tipo de condiciones las respuestas se van a hacer más específica y va a poder filtrar las cambiantes de registro.

Aplicando las redes bayesianas se puede dar soporte para la toma de decisión y así reducir uno de los problemas más comunes al momento de un registro en paralelo, como podría serlo en registro vacíos donde no hay datos de donde buscar. (Bonilla, Ojeda y Echevarría, 2006)

b) Autor: Elena Durán, Rosanna Costaguta

Título: APLICACIÓN DE TÉCNICAS DE MINERÍA DE DATOS PARA DESCUBRIR ESTILOS DE APRENDIZAJE, 2007.

Correlación:

Los alumnos aprenden de muchas maneras: observando y oyendo, reflexionando y actuando, razonando lógica e intuitivamente, memorizando y visualizando, creando analogías y modelos matemáticos. Además, los procedimientos de lección son diversos. Cuanto aprende un estudiante es dependiente de su capacidad innata y de su elaboración previa, pero además de la compatibilidad entre su estilo de estudio y el estilo de lección del instructor. Como docentes de la carrera Licenciatura en Sistemas de Información de la Facultad Nacional de Santiago del Estero, pensamos que resultaría servible saber las propiedades del perfil de estudio de nuestros alumnos para adecuar las tácticas de lección. Con este propósito encuestamos a ochenta y nueve alumnos, aplicando el Test de Felder y Soloman.

La información obtenida originó una base de datos de información a la que se le aplicó el Desarrollo de Hallazgo de Conocimiento (Knowledge Discovery from Database - KDD). De esta forma determinamos que hay un prominente nivel de homogeneidad en el estilo de estudio de los estudiantes. La identificación del cluster dominante permitió comprender además las propiedades del estilo de estudio compartido por la mayor parte de los estudiantes. Sabiendo la información descubierta se proponen tácticas de participación didáctica y de exhibición de material educativo acorde al estilo dominante. (Durán, Costaguta, 2007)

c) Autor: José C. Riquelme, Roberto Ruiz, Karina Gilbert

Título: MINERÍA DE DATOS: CONCEPTOS Y TENDENCIAS, 2006.

Correlación:

En la actualidad, la minería de datos (MD) está consiguiendo cada vez más capturar la atención de las compañías. Por el momento es infrecuente oír oraciones como “deberíamos dividir a nuestros usuarios usando utilidades de MD”, “la MD incrementará el agrado del cliente”, o “la rivalidad está usando MD para ganar cuota de mercado”. Por otro lado, todo apunta a que más temprano que tarde la minería de datos va a ser utilizado por la sociedad, por lo menos con el mismo peso que en la actualidad tiene la Estadística.

Así que ¿qué es la minería de datos y qué beneficios aporta? ¿Cómo puede influir esta tecnología en la resolución de los problemas diarios de las empresas y la sociedad en general? ¿Qué tecnologías están detrás de la minería de datos? ¿Cuál es el ciclo de vida de un proyecto típico de minería de datos? En este artículo, se intentarán aclarar estas cuestiones mediante una introducción a la minería de datos: definición, ejemplificar problemas que se pueden resolver con minería de datos, las tareas de la minería de datos, técnicas usadas y finalmente retos y tendencias en minería de datos. (Riquelme, Ruiz y Gilbert, 2006)

d) Autor: Yasunari del Valle Ramírez León, José Antonio Ortega Carrillo

Título: DIAGNÓSTICO DEL ESTILO DE APRENDIZAJE PREDOMINANTE BASADO EN MINERÍA DE DATOS EL MODELO DE FELDER: APLICACIONES AL ELEARNIG 3.0., 2012.

Correlación:

En este texto se detalla el desarrollo de diagnóstico del estilo de estudio de un grupo de alumnos concreto, empleando el Cuestionario sobre el Índice de Estilos de Estudio de Felder y minería de datos, por medio de la aplicación del análisis apoyado en clustering.

En inicio, en forma de contextualización se enfoca lo referido al tema teórico del modelo de estilos de estudio de Felder y el Cuestionario sobre el Índice de Estilos de Estudio de Felder, los conceptos relativos a la minería de datos aplicada a la educación y una corto sinopsis del programa WEKA. Más adelante, se expone la

correspondencia entre los estilos de estudio y las probables respuestas del Cuestionario sobre el Índice de Estilos de Estudio, empleado para saber el estilo de estudio de la gente estudiada. De la misma manera, se muestra el desarrollo de obtención del estilo de estudio predominante en la exhibe poblacional en estudio, según con el análisis apoyado en los clústeres obtenidos. Por último, se dan a conocer las conclusiones propias de la materia abordada. (Ramírez y Ortega, 2012)

e) Autor: José A. Martí Arias

Título: ESTILOS DE APRENDIZAJE Y MODELO DE USUARIO EN LOS SISTEMAS HIPERMEDIA ADAPTATIVOS, 2015.

Correlación:

Este trabajo muestra una introducción al inconveniente de los sistemas de lección automatizados que se fundamentan en la amoldación del sistema a los individuos potenciales. En especial los experimentos se hacen con alumnos de la Carrera de Microbiología de la Facultad de la Habana. Se parte de la determinación de los distintos estilos de estudio de los alumnos, usando para eso diferentes tests (Felder-Silverman, Predominancia Hemisférica.

Cuadrantes Cerebrales y Canal perceptual predominante). Se muestran los resultados de los 4 primeros tests aplicados y alumnos universitarios. Además, se analizan las propiedades simples resultantes de otros destacables tests como el de Honey-Alonso y Memletics, con la intención de incorporarlos además al modelo de usuario primordial dentro de un sistema adaptativo para la educación. En una segunda etapa se determinan las propiedades y estilos de lección de los instructores comprometidos en la docencia de cada curso con la intención de usar la pedagogía correcta, oséa, saber y adaptar las más utilizadas opciones de lección y estudio de manera que tomen presente, tanto las propiedades del instructor como los estilos de estudio de sus alumnos con el propósito de conseguir un mejor desempeño académico. La última fase es la concerniente a la implementación de un sistema ó módulo adaptativo, que esté basado en el modelo de usuario el cual incluye los perfiles del usuario previamente determinados. (Martí, 2015)

2.2. MARCO TEORICO

2.2.1. KDD

KDD es la extracción no trivial de información probablemente servible desde un enorme volumen de datos, en el cual la información está implícita, donde hablamos de interpretar enormes proporciones de datos y hallar relaciones o patrones como entre otras cosas hallar cuantos alumnos fueron a clases en los exámenes finales y parciales. (Webmining Consultores, 2011)

A.1. Características del KDD

Es procesar de forma automática enormes proporciones de datos para hallar conocimiento servible para un usuario y satisfacerle sus misiones, es el propósito primordial del sector de Hallazgo de Conocimiento en Bases de Datos o KDD (Knowledge Discovery from Data base). Este es el campo que está evolucionando para proveer resoluciones al análisis automatizado.

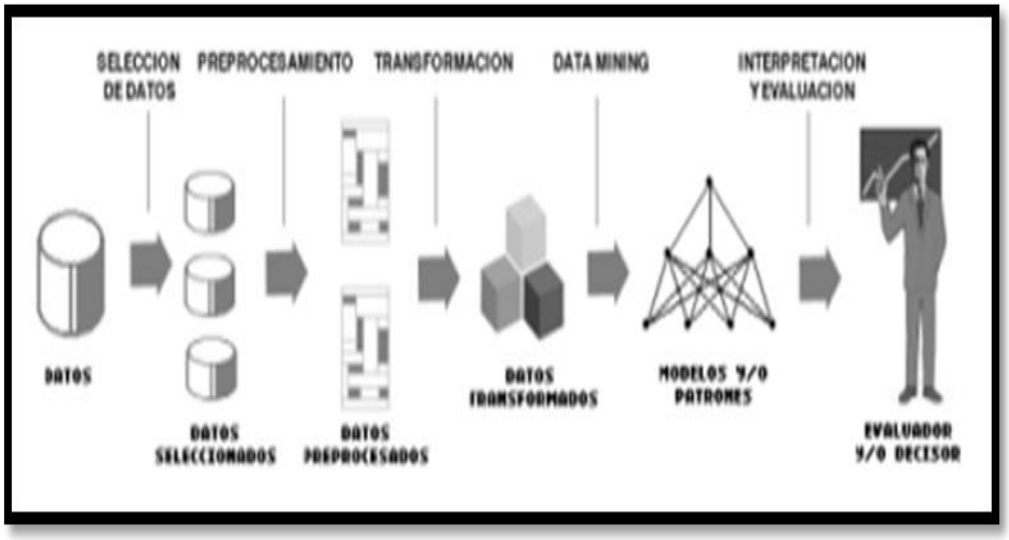


Figura 4. Proceso del KDD.
Adaptado de “Proceso del KDD” por Webmining Consultores, 2011

A.2.1. Selección de Datos

Tiene como funcionalidad la selección esos datos o exhibe de datos que sean servibles y necesarios sobre los cuales se realizara el hallazgo. La selección puede cubrir tablas, atributos y datos, con una correcta selección de atributos y datos, se va a mejorar el desarrollo de los algoritmos a usar, además generan más acelerado los resultados y posibilita el entendimiento del hallazgo del conocimiento obtenido.

Antes de tomar algún elección en relación a la selección de datos o atributos es requisito tener conocimiento sobre lo que quiere investigar, debido a que las elecciones que se tomen sobre los atributos afectan de manera directa a los resultados, y es posible que se seleccionen más atributos, pero esto no significa que se va conseguir un mejor resultado, de esta forma se tienen la posibilidad de añadir, remover o cambiar atributo que se utilicen a lo largo de todo el desarrollo hasta que estos empiecen a crear superiores resultados y sucede que mientras mejor se analicen los datos a usar los resultados tienen la posibilidad de ser más correctos.

A.2.2. Preprocesamiento

En esta etapa del desarrollo, se ejecuta el pre procesamiento y limpieza de los datos, que radica en remover datos incoherentes, registros repetidos y basura en los datos producidos por las migraciones que se desarrollan en ellos, y si es requisito y correspondiente se eliminan datos que no sean servibles en el desarrollo, pero es habitual hallar datos que tienen dentro errores, pero en esta situación conveniente sostener los datos.

El pre procesamiento de los datos, es la etapa donde se ocupa la más grande parte del tiempo para investigar y elaborar los datos y por eso mismo esto es primordial dado que de la elaboración de los datos es dependiente mayormente de los resultados.

La información que se va formando del desarrollo de limpieza y el pre procesamiento de los datos nos van provocando información o datos muy importantes que pasando cada desarrollo los resultados esperados nos llevaran a tomar increíbles elecciones.

A.2.3. Transformación

La transformación de los datos se apoya en que es requisito que los datos sean enriquecidos con otras fuentes de información ya sean internas o externos para achicar los datos y el número de cambiantes y transfórmalos en un formato que ayude el mejoramiento de la información para llegar al conocimiento requerido.

La transformación de los datos es primordial dado que hasta aquí es donde se toma el formato requerido para enlazarlo a la minería de datos y obtener el saber que hay que tomar al filtrar los datos que se van consiguiendo por medio del desarrollo. En

este punto se tienen dentro operaciones simples sobre datos, en el cual consisten en el filtrado de los datos para reducirlos y achicar qué llevar a cabo con los datos que logren ser faltantes, una de las tareas o puntos que puede cubrir el pre procesamiento es:

- Derivar nuevos atributos: que consiste en crear nuevos campos explícitos con relaciones entre atributos conocidos que puede ser el análisis más sencillo.
- Agrupamiento: este consiste en donde hay relaciones uno a muchos y podemos convertir estas relaciones en uno a uno para mejorar un mejor análisis y encontrar un mejor resultado.

A.2.4. Data Mining

La etapa más completa del desarrollo del KDD, por medio de el cual se extraen los patrones según las misiones proposiciones dentro de la exploración, minería de datos y KDD se identifican como si fuera uno mismo.

Como ya se comentó, Minería de datos (una etapa dentro del desarrollo terminado de KDD) intenta conseguir patrones o modelos desde los datos recopilados. Elegir si los modelos obtenidos son servibles o no acostumbra necesitar una valoración subjetiva por parte del usuario. Los algoritmos de Minería de datos acostumbran tener tres elementos.

- El modelo, que contiene parámetros que han de fijarse a partir de los datos de entrada.
- El criterio de preferencia, que sirve para comparar modelos alternativos.
- El algoritmo de búsqueda (como cualquier otro programa de IA).
- La etapa Minería de Datos, abarca extracción del conocimiento en bases de datos donde los datos son fuentes importantes ya que cada dato se va analizando por los diferentes procesos antes mencionados modelo, criterio y algoritmo y estos datos nos llevan obtener mejores resultados.

El ingrediente Minería de datos de KDD envuelven la aplicación repetitiva de los procedimientos que usa para la extracción de los datos y atributos dado que es sustancial realizar una cantidad enorme de algoritmo con el objetivo de llegar a un mismo resultado esperado, más allá de que se tenga una misión clara o definida no está demás buscar otras misiones dado que esto nos puede contribuir a conseguir superiores en el análisis y tomar superiores elecciones.

A.2.5. Interpretación y Evaluación

Una vez obtenido el modelo, hay que seguir a su validación comprobando que las conclusiones que arroja son válidas y bastante satisfactorias. En la situación de haber obtenido numerosos modelos por medio de la utilización de diferentes técnicas, se tienen que contrastar los modelos en busca de aquel que se ajuste mejor al inconveniente. Si ninguno de los modelos consigue los resultados esperados, debe alterarse alguno de los pasos anteriores para crear nuevos modelos. Si el modelo final no superara esta evaluación el desarrollo se podría reiterar desde el inicio o, si el profesional lo considera oportuno, desde alguno de los pasos anteriores. Esta retroalimentación se va a poder reiterar cuantas ocasiones se considere primordial hasta conseguir un modelo válido.

Una vez obtenido el modelo, hay que seguir a su validación comprobando que las conclusiones que arroja son válidas y bastante satisfactorias. En la situación de haber obtenido numerosos modelos por medio de la utilización de diferentes técnicas, se tienen que contrastar los modelos en busca de aquel que se ajuste mejor al inconveniente. Si ninguno de los modelos consigue los resultados esperados, debe alterarse alguno de los pasos anteriores para crear nuevos modelos. Si el modelo final no superara esta evaluación el desarrollo se podría reiterar desde el inicio o, si el profesional lo considera oportuno, desde alguno de los pasos anteriores. Esta retroalimentación se va a poder reiterar cuantas ocasiones se considere primordial hasta conseguir un modelo válido

Comúnmente, las técnicas de minería de datos se aplicaban sobre información contenida en almacenes de datos. De hecho, muchas enormes compañías e instituciones han desarrollado y alimentan bases de datos principalmente diseñadas para proyectos de minería de datos en las que centralizan información probablemente servible de todas sus superficies de negocio.

A.3. Minería de datos

La Minería de Datos posibilita la extracción de información importante de enormes bases de datos, información que exhibe conocimiento por medio de patrones ocultos, tendencias y correlaciones que indican conjeturas para solucionar inconvenientes del negocio, ofreciendo una virtud competitiva, estos patrones van a ser seleccionados por el responsable del sector de sistemas con el objetivo de poder tomar elecciones asertivas.

A.3.1. Característica de Minería de Datos

El propósito de la Minería de Datos, es utilizar los datos accesibles para crear un modelo que detalla una variable en especial en relación a los otros datos accesibles. Con la Minería de Datos auxiliará al responsable del sector de sistemas en el procesamiento de vastas reservas de datos para conocer “relaciones insospechadas”, entre otras cosas, porcentajes de alumnos que postularon; en ese caso el sector administrativo es el solicitado de tener esos ficheros.

A.3.2. Algoritmos de Minería de Datos

Los algoritmos de minería de datos se clasifican en dos grandes categorías:

- Supervisados o predictivos.
- No supervisados o de descubrimiento del conocimiento.

El propósito de la Minería de Datos, es utilizar los datos accesibles para crear un modelo que detalla una variable en especial en relación a los otros datos accesibles. Con la Minería de Datos auxiliará al responsable del sector de sistemas en el procesamiento de vastas reservas de datos para conocer “relaciones insospechadas”, entre otras cosas, porcentajes de alumnos que postularon; en ese caso el sector administrativo es el solicitado de tener esos ficheros.

Tabla 8
Fases del Algoritmo en la Minería de Datos

PROCEDIMIENTO	COMPONENTE
Realizar Búsqueda	Minería de datos
Definir lugar de búsqueda	Base de datos
Establecer criterios de búsqueda	Población
Concretar método de búsqueda	Toma de decisiones
Guardar lo encontrado	Técnica de procesamiento y obtención de conocimiento

Adaptado de “Proceso del KDD” por Webmining Consultores, 2011

El sistema construido tiene los siguientes componentes funcionales: los algoritmos o programas que buscan (Técnicas); en dónde se busca (la base de datos); qué se busca (tipificación de lo interesante); cómo se busca (orden, secuencia de búsqueda).

2.2.2. Modelo de Estilos de Aprendizaje

Un modelo de estilos de estudio clasifica a los alumnos según la forma en cómo ellos reciben y procesan la información. El modelo propuesto por Felder y Silverman (1988):

> Es particularmente aplicable a estudiantes de ingeniería. Estos autores proponen, paralelamente, un modelo de estilo de enseñanza que clasifica los métodos instruccionales de acuerdo a cuan bien direccionan los componentes del estilo de aprendizaje propuesto. En su trabajo original, los autores consideraron cinco dimensiones de análisis: percepción, entrada, organización, procesamiento y comprensión; pero en una posterior versión suprimen la dimensión de organización (inductivo y deductivo). (parr.1).

A.1. APRENDIZAJE VISUAL Y VERBAL

Los estudiantes visuales recuerdan mejor lo que ven: figuras, diagramas, cuadros, líneas de tiempo, películas, demostraciones. Los estudiantes auditivos recuerdan más lo que escuchan y mucho más lo que ellos dicen. Ellos aprenden a partir de la discusión y prefieren las explicaciones verbales a las demostraciones visuales. Aprenden efectivamente cuando ellos pueden explicarles a otros.

A.2. APRENDIZAJE ACTIVO Y REFLEXIVO

El desarrollo mental complejo por el cual la información percibida es transformada en conocimiento puede ser convenientemente agrupado en dos categorías: experimentación activa y observación reflexiva. La primera supone llevar a cabo algo en el planeta de afuera con la información (discutirla, explicarla o chequearla de alguna manera), y la segunda supone investigar y manejar la información introspectivamente. Según esto, un estudiante activo es el que se siente más práctico con la experimentación activa que con la observación reflexiva, al revés de un estudiante reflexivo. Los alumnos activos no aprenden bastante en ocasiones en las que ellos tienen que estar pasivos, así como lecturas; trabajan bien en grupos y por lo general son experimentalistas. Los alumnos reflexivos no aprenden bien en ocasiones que no les ofrecen la posibilidad de reflexionar sobre la información que se les presenta; trabajan mejor solos o a lo sumo con un individuo más; por lo general son teóricos.

A.3. APRENDIZAJE SECUENCIAL Y GLOBAL

El estudio secuencial es el aplicado en la mayor parte de la educación formal, e supone la exhibición de material en un orden de progresión lógica, con el avance del estudio regido por el tiempo y el calendario. Cuando un cuerpo de material fué cubierto los alumnos son evaluados y recién tienen la posibilidad de pasar al nivel o cuerpo siguiente. El estudio global es aquel que no se funciona por el tiempo ni el calendario, por el opuesto, se tienen la posibilidad de pasar días o semanas ocupados en solucionar un fácil inconveniente o demostrando una comprensión rudimentaria hasta que de repente se le “encienden las luces” y logran una ligera comprensión del todo.

2.2.3. Instrumentos o Test Aplicados

a) Con el objetivo de reducir el número de preguntas del **cuestionario ILS (Index of Learning Styles)** y generar un cuestionario adaptativo se ha llevado a cabo una investigación en la Univ. Autónoma del Perú.

Tabla 9
Dimensiones de estilos de aprendizaje de Felder
Adaptado de “Estilos de Aprendizaje y Modelo de Usuario en los Sistemas Hipermedia Adaptativos” de Marti, 2015

Definiciones	Dimensiones		Definiciones
Hace	**Activo**	**Reflexivo**	Piensa
Aprende hechos	**Sensitivo**	**Intuitivo**	Aprende conceptos
Requiere dibujos	**Visual**	**Verbal**	Requiere leer o disertar
Paso a paso	**Secuencial**	**Global**	Marco general
Deriva principios de hechos	**Inductivo**	**Deductivo**	Deriva resultados de principios

El cuestionario de Felder-Silverman terminado radica de 44 cuestiones. Aquí debe seleccionarse la alternativa (a) ó (b) para señalar su respuesta. En la situacion de que (a) como (b) simulen ser sus respuestas, se selecciona aquella que se coloque más muchas veces. Felder y Soloman [Felder y Soloman, 2004] sugieren agrupar a los alumnos en cinco categorías para todas las dimensiones. ILS hace 11 cuestiones para todas las dimensiones del modelo y la puntuación obtenida por cada estudiante se calcula restando las respuestas “b" a las respuestas “a".

Tabla 10

Hoja de calificación del cuestionario de Felder (ILS)

Adaptado de "Estilos de Aprendizaje y Modelo de Usuario en los Sistemas Hipermedia Adaptativos" Marti, 2015

Activo/Reflexivo			**Sensitivo/Intuitivo**			**Visual/Verbal**			**Secuencial/Global**		
Quiz	a	b	Quiz	a	b	Quiz	a	b	Quiz	a	b
1			2			3			4		
5			6			7			8		
9			10			11			12		
13			14			15			16		
17			18			19			20		
21			22			23			24		
25			26			27			28		
29			30			31			32		
33			34			35			36		
37			38			39			40		
41			42			43			44		
Total			**Total**			**Total**			**Total**		
%			%			%			%		

Así el resultado final del cuestionario son 4 puntuaciones (números impares entre 11 y -11), una para cada extensión. Hay, por consiguiente, 12 probables puntuaciones para cada una. El ejemplo siguiente exhibe el papel de resultados de la evaluación del Cuestionario ILS.

- **La dimensión activa/reflexivo**

  Está relacionada históricamente con la utilización de simulaciones en los tutoriales y con el trabajo en grupo. Esta extensión se usará para la generación automática de grupos.

- **La dimensión sensorial/intuitivo**

  Está relacionada con el contenido exacto que se expone. Los estudiantes sensoriales eligen la exhibición de las explicaciones después de los ejemplos y al revés para los estudiantes intuitivos. Esto piensa, por una parte, la necesidad de una relación de pertenencia entre explicaciones y ejemplos, y por otra, el problema de probables conflictos entre la secuenciación establecida por la extensión secuencial/global y la establecida por esta extensión.

- **La dimensión visual/verbal**

  Es una extensión difícil de adaptar en la educación a través de Internet ya que en tanto que en las dimensiones secuencial/global y sensorial/intuitivo el sistema ajusta la secuenciación del curso, la adaptación a la extensión visual/verbal va a poder

ocasionar la supresión de material educativo. De hecho, si pudiésemos expresar el mismo contenido de ambas formas, a través de expresiones y a través de imágenes (algo difícil de asegurar a priori), no logramos hallar ninguna razón para no enseñar ambas explicaciones del mismo método a todos los estudiantes. A lo destacado la integración o no de imágenes va a poder depender en su mayoría de los requerimientos técnicos.

- **La dimensión secuencial/global**

Perjudica a la visión del estudiante, va a ser el punto desde donde el estudiante mira el curso. Los alumnos secuenciales eligen ver el curso desde dentro, mientras los alumnos globales eligen observarlo desde fuera. Enseñar el curso de la forma deseada por cada uno va a poder arreglar el inconveniente de estar perdido en el hiperespacio para los alumnos secuenciales y el inconveniente de la perdida de independencia para los alumnos globales.

Tabla 11
Perfil de un estudiante modelo de Felder

Adaptado de "Estilos de Aprendizaje y Modelo de Usuario en los Sistemas Hipermedia Adaptativos" de Martí, 2015

	11	9	7	5	3	1	1	3	5	7	9	11	
ACTIVO									X				REFLEXIVO
SENSORIAL											X		INTUITIVO
VISUAL								X					VERBAL
SECUENCIAL									X				GLOBAL

- Si el puntaje en la escala está **entre 1 - 3** la persona presenta un **equilibrio apropiado** entre los dos extremos de esa escala. El estudiante *TIENE UN EQUILIBRIO ENTRE VISUAL Y VERBAL.*
- Si el puntaje está **entre 5 - 7** presenta una **preferencia moderada** hacia uno de los dos extremos de la escala y aprenderá más fácilmente si se le brinda apoyo en esa dirección. El alumno es *MÁS REFLEXIVO QUE ACTIVO Y MÁS GLOBAL QUE SECUENCIAL.*
- Si el puntaje en la escala está **entre 9 - 11** presenta una **preferencia muy fuerte** por uno de los dos extremos de la escala. Usted puede llegar a presentar dificultades para aprender en un ambiente en el cual no cuente con apoyo en esa dirección. El estudiante *ES MUCHO MÁS INTUITIVO QUE SENSITIVO.*

b) Predominancia hemisférica

Aunque cada individuo usa por siempre todo su cerebro, hay relaciones continuas entre los dos hemisferios, y por lo general uno es más activo que el otro. Cada hemisferio es el responsable de medio cuerpo ubicada en el lado opuesto:
ósea, el hemisferio derecho dirige la sección izquierda del cuerpo, en tanto que el hemisferio izquierdo dirige la sección derecha.

En la determinación de la dominancia de los hemisferios influyen componentes sociales. Cada hemisferio trata la información que recibe de diferente forma, ósea, hay diferentes maneras de pensamiento asociadas con cada hemisferio.

- ✓ Hemisferio Izquierdo del cerebro está más especializado en el manejo de los símbolos de algún tipo: lenguaje, álgebra, símbolos químicos, partituras musicales. Es más analítico y lineal y procede de manera lógica.
- ✓ El hemisferio derecho es más efectivo en la percepción del espacio, es más global, sintético e intuitivo. Es imaginativo y emocional.

c) Cuadrantes cerebrales predominantes

Los cuatro cuadrantes representan cuatro formas distintas de operar, de pensar, de crear, de aprender y, en suma, de convivir con el mundo.
Las características generales de los diferentes cuadrantes cerebrales son:

✓ Cortical izquierdo. - Alumnos que tienen necesidad de hechos y dan prioridad al contenido. Les agradan las clases sólidas, argumentadas, apoyadas en los hechos y las pruebas. Van a clase a estudiar, tomar apuntes, seguir en el software para conocerlo bien en el final del curso. Son buenos estudiantes a condición de que se les dé materia.

✓ Cortical derecho. - Requieren apertura y perspectiva de futuro a la larga. Son intuitivos y animosos. Toman escasas notas de clase porque sabe elegir lo fundamental. Algunas veces llaman la atención como soñadores, o de estar desconectad les, pero otras sorprenden con visualizaciones inesperadas y proyectos auténticos.

✓ Límbico izquierdo. - Se atienen a la forma y a la organización. Son metódicos, organizados, y con continuidad meticulosos; los desborda la toma de apuntes porque intentan ser claros y limpios. Llegan a copiar otra vez un cuaderno o una

asignatura por encontrarlo raro o sucio. Les agrada que la clase se desarrolle según un procedimiento popular y rutinario.

- ✓ Límbico derecho. - Se atienen a la comunicación y a la relación. Trabajan por el sentimiento e instinto. Aprecian las pequeñas astucias de la pedagogía. Trabajan si el instructor es de su agrado; se bloquean y despistan de forma sencilla si no se piensan sus progresos o adversidades. No soportan críticas severas. Le agradan algunas materias, detestan otras y lo demuestran. Aprecian las salidas de clase, juegos, videos y todo aquello que no se parezca a una clase.

d) Determinación del canal perceptual predominante.

El test sirve para la determinación del canal perceptual dominante, apoyado en el Modelo de la Programación Neurolingüística (PNL). Se cree que un 40% de la gente es visual, un 30% auditiva y un 30% kinestésica. Por otro lado, el hombre es con la capacidad de retener un 20% de lo que escucha; un 50% de lo que ve y escucha, y un 75% de lo que ve, escucha y costumbre. Debemos recordar un proverbio chino que reza como: “DÍGAME Y OLVIDO, MUÉSTREME Y RECUERDO, INVOLÚCREME Y COMPRENDO”

Para la PNL, el canal perceptual por donde nos apropiamos de los contenidos (el ojo, el escuchado, o el cuerpo en general) resulta primordial en nuestras opciones de estudio. La mayor parte de nosotros usamos estos canales en forma desigual, potenciando unos y desatendiendo otros. Esto otorga origen a tres estilos básicos de aprendizaje: visual, auditivo y kinestésico. Entre otras cosas, si intentamos suponer en una localidad que hayamos visitado una vez, algunos recordaremos mejor el aspecto de sus inmuebles (visual), otros su nombre (auditivo), otros el calor o frío que hacía cuando estuvimos ahí (kinestésico).

La persona kinestésica es muy sentimental, sensitiva y emocional. Demuestran su sensibilidad y manifiestan espontáneamente sus sentimientos. Se relacionan muy de forma sencilla con otra gente. El aspecto no les atrae bastante, algunas ocasiones su forma de vestir tiende a ser descuidada y puede no unir sólo les importa sentirse prácticos. Se mueven bastante, pero con soltura y simplicidad. Sus posiciones son muy relajadas, con los hombros bajos y caídos. Sus movimientos son pausados y tranquilos. Gesticulan bastante, se tocan y tocan todo el tiempo a los otros.

La persona auditiva tiende a ser más sedentaria que la visual. Es más cerebral que otros y tiene mucha vida interior. Está muy entusiasmado en escuchar. La persona auditiva es increíble conversadora y tiene una capacidad de ordenar mentalmente sus ideas. Algunas veces se ve estar de mal humor gracias a su sensibilidad a algunos tipos de ruidos. Comúnmente son muy serios y no sonríen bastante. Su forma de vestir jamás será tan sustancial como sus ideas. Su estilo tiende a ser conservador y elegante.

La persona visual sabe el planeta de la misma forma que lo ve y el aspecto de las cosas es lo más relevante. Cuando recuerda algo lo realiza con apariencia de imágenes; transforma las expresiones en imágenes y cuando piensa algo del futuro lo visualiza. Son muy organizados, les gusta mucho ver el planeta ordenado y limpio, siempre están controlando las cosas para asegurarse de que están bien ubicadas. La multitud visual tiende a ser esbelta. Su posición es algo rígida, con la cabeza inclinada hacia enfrente y los hombros en prominente. Se muestra bien vestida y siempre se le ve arreglada y limpia. El aspecto le es primordial, combina bien su ropa y la escoge con precaución.

e) Instrumento de Alonso, Gallego y Honey (1995)

Las aportaciones y vivencias de Honey y Mumford fueron agarradas en 1992 en España por Catalina Alonso, quien adaptó el cuestionario LSQ de Estilos de Estudio al tema estudiantil y al idioma español. Este cuestionario ajustado es el CHAEA (Cuestionario Honey-Alonso de Estilos de Aprendizaje) y no es un test de sabiduría, ni de personalidad. El cuestionario CHAEA tiene 80 ítems, cada ítem se responde con un signo (+) sí se está en concordancia y con un (-) sí se no está de acuerdo. (Alonso, Gallego y Honey, 1995)

Los resultados del cuestionario se plasman en una hoja que sirve para determinar las preferencias en cuanto a los Estilos de Aprendizaje:

- Activo.
- Reflexivo.
- Teórico.
- Pragmático.

Todos los ítems están organizados aleatoriamente formando un solo grupo. La puntuación absoluta que el sujeto obtiene en cada grupo de 20 ítems, va a ser el nivel que alcance en todos los 4 estilos de estudio.

f) También es el de S. Whiteley y Whiteley (2003).

Los esposos Whiteley desarrollaron el proyecto **Memletics** de aprendizaje acelerado que incluye un cuestionario de estilos de aprendizaje con 70 preguntas que se tienen que evaluar con las escalas 0, 1 y 2 (así no soy, así parcialmente soy y así soy). Las 70 preguntas corresponden a 7 diferentes estilos de aprendizaje: visual-espacial, auditivo-aural, verbal-lingüístico, físico-kinestésico, lógico-matemático, social-interpersonal y solitario intrapersonal. (García, 2009)

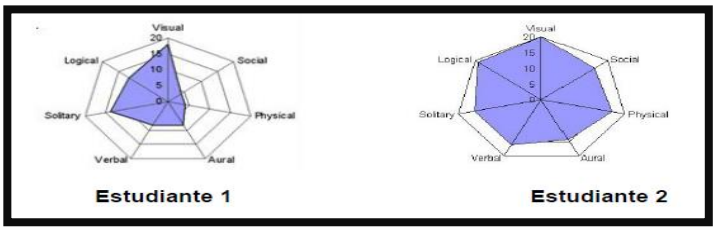


Figura 5. Resultados de MEMLETICS
Adaptado de "Instrumento de medición de aprendizaje" por García, 2009

g) Los Sistemas Hipermedia Adaptativos

Los sistemas hipermedia adaptativos (SHA) son fundamentalmente sistemas hipermedia con aptitud de adaptar la exhibición de contenidos, los formatos de información y las configuraciones de navegación a las propiedades de sus potenciales individuos, esto con el objetivo de proveer la información más importante y en los formatos más correctos para todos ellos. Particularidades como el nivel de entendimientos, las vivencias anteriores, el idioma y los intereses fueron al principio modeladas para producir la amoldación en los primeros SHA premeditados a beneficiar el estudio, en un sector específica del conocimiento. Más adelante, estos tipos de sistemas han incorporado otras cambiantes, los estilos cognitivos o los estilos de estudio de los estudiantes-usuarios. Esta novedosa forma de basar la amoldación se fundamenta en que diferentes alumnos no logran un igual nivel de desarrollo frente a una exclusiva forma de lección, que es la que fundamentalmente ofrecen los SH, sino que necesitan de tácticas instruccionales correctas a la forma como procesan y perciben la información.

Los SHA académicos desarrollados y que basan la adaptación en los estilos de estudio de sus potenciales individuos, se han sustentado en diferentes modelos o

teorías de estilos de estudio. En algunos sistemas la adaptación se materializa de forma exclusiva en la exhibición de los contenidos, en otros, se integran además recomendaciones para las configuraciones de navegación. Diferentes y variadas tácticas instruccionales, que tienen dentro diferentes formatos de información, organización de contenidos y secuencias de actividades, son utilizadas para ajustar los sistemas a las necesidades de los estudiantes-usuarios en función de sus respectivos estilos de aprendizaje.

h) Algunas estrategias de adaptación

Para cambiar los contenidos y la estructura de un curso a las características del estudiante mientras aprende de forma individual se emplean las denominadas tácticas de adaptación que admiten:

- ✓ **Modificación de la estructura del curso**: una labor cierta puede ser presentada a diferentes alumnos en diferentes puntos del curso, ósea, en diferentes instantes en relación al avance del curso. Esto otorga lugar a construcciones del curso personalizadas para cada estudiante.
- ✓ **Presentación o no de ciertas tareas**: una labor específica puede ser presentada a un estudiante mientras ejecuta un curso, mientras esa misma labor puede ser omitida para otro estudiante que ejecuta el mismo curso.
- ✓ **Requisitos previos**: en un sistema adaptativo dado un grupo de tareas, se tienen la posibilidad de detallar dependencias entre ellas, de forma que si no se hizo alguna labor o tareas antes no se logre entrar a otras. Se tienen la posibilidad de detallar diferentes prerrequisitos para diferentes alumnos en funcionalidad de sus propiedades personales y de sus acciones anteriores.
- ✓ **Adaptación de la secuenciación de contenidos**: una labor puede dividirse en subtareas, y se puede entablar el orden de esas subtareas por medio de reglas de secuenciación. Es viable entablar un orden concreto para esas subtareas (AND), dejar a los alumnos seleccionar el orden de ejecución (ANY), o dejarles seleccionar entre diferentes subtareas con el mismo propósito (OR). Por una sección, diferentes reglas de secuenciación tienen la posibilidad de ser establecidas para diversos tipos de alumnos, y por otra, el orden entre las distintas subtareas de una labor con regla de secuenciación AND, puede además ser diferente en relación del tipo de estudiante.

i) Adaptación de contenidos

Las páginas mostradas a los alumnos por el sistema se desarrollan de manera dinámica desde los extractos HTML. Variantes de esos extractos, introduciendo diferentes elementos multimedia, tienen la posibilidad de ser dentro en la página que se expone al estudiante para la ejecución de una cierta labor, de manera que la aclaración, el ejemplo o el enunciado del ejercicio más correspondiente para cada estudiante puede ser constituido en tiempo de ejecución. Los distintos extractos tienen la posibilidad de hallarse en diferentes lenguajes. Una relación entre los estilos de estudio del modelo de Felder-Silverman y los aspectos que tienen la posibilidad de ser adaptados en un sistema de lección por medio de Internet.

Tabla 12
Adaptación de Contenidos
Adaptado de "Estilos de aprendizaje y de enseñanza en la educación de ingeniería" por Felder y Silverman, 1988

¿Qué podemos adaptar?	Activo-Reflexivo	Sensorial-Intuitivo	Visual-Verbal	Secuencial-Global
	Participación	Contenido	Presentación	Perspectiva

2.2.4. CRISP-DM

CRISP-DM, que son las siglas de Cross-Industry Standard Process for Data Mining, es un método probado para orientar sus trabajos de minería de datos.

- Como metodología, incluye descripciones de las fases normales de un proyecto, las tareas necesarias en cada fase y una explicación de las relaciones entre las tareas.
- Como modelo de proceso, CRISP-DM ofrece un resumen del ciclo vital de minería de datos.

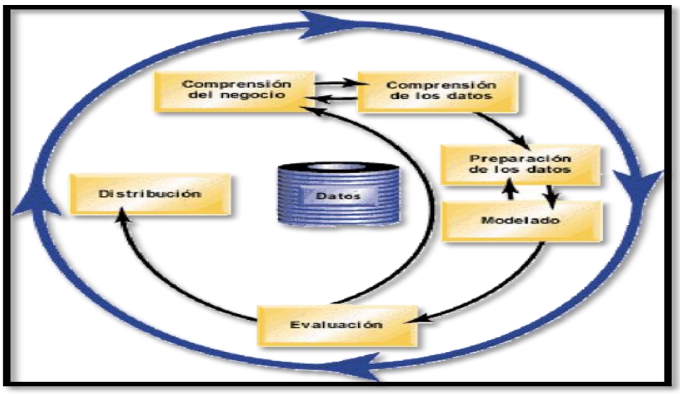


Figura 6. Ciclo de Vida de CRISP-DM
Adaptado de "CRISP-DM" por IBM, 2012

El ciclo vital del modelo contiene seis fases con flechas que indican las dependencias más importantes y frecuentes entre fases. La secuencia de las fases no es estricta. De hecho, la mayoría de los proyectos avanzan y retroceden entre fases si es necesario.

El modelo de CRISP-DM es maleable y se tienen la posibilidad de ajustar de forma sencilla. Entre otras cosas, si su organización intenta descubrir ocupaciones de blanqueo de dinero, posiblemente requiera hacer una criba de enormes proporciones de datos sin un propósito de modelado concreto. En vez de hacer el modelado, su trabajo se centrará en examinar y ver datos para conocer patrones sospechosos en datos financieros. CRISP-DM facilita hacer un modelo de minería de datos que se ajuste a las necesidades concretas.

En tal circunstancia, las fases de modelado, evaluación e utilización tienen la posibilidad de ser menos importantes que las fases de elaboración y comprensión de datos. No obstante, es primordial tener en cuenta algunas cuestiones que emergen a lo largo de fases posteriores para la idealización a la larga y objetivos futuros de minería de datos.

A.1. Fase de comprensión del negocio o problema

La primera etapa de la guía de referencia CRISP-DM, llamada etapa de comprensión del negocio o inconveniente, es posiblemente la más relevante y aglutina las tareas de comprensión de los objetivos y requisitos del emprendimiento desde una visión empresarial o institucional, con el objetivo de convertirlos en objetivos técnicos y en un plan de emprendimiento. Sin conseguir abarcar estos objetivos, ningún algoritmo por muy exclusivo que sea, dejará conseguir resultados fiables. Para conseguir el más destacable beneficio de Data Mining, es requisito comprender de la forma más completa el inconveniente que se quiere solucionar, esto dejará juntar los datos correctos e interpretar de manera correcta los resultados. En esta etapa, es primordial la aptitud de poder transformar el saber conseguido del negocio, en un inconveniente de Data Mining y en un plan preliminar cuya misión sea el lograr los objetivos del negocio. Una especificación de todas las primordiales tareas que conforman esta etapa es la siguiente:

Saber los objetivos del negocio. Esta es la primera labor a desarrollar y tiene como misiones, saber cuál es el inconveniente que se quiere solucionar, por qué la necesidad de usar Data Mining y determinar los criterios de triunfo. Los inconvenientes tienen la posibilidad de ser distintos como, entre otras cosas, descubrir estafa en la utilización de tarjetas de crédito, descubrimiento de intentos de ingreso indebido a un sistema, garantizar el triunfo de una cierta operación publicitaria. En relación a los criterios de triunfo, estos tienen la posibilidad de ser de tipo cualitativo, en cuyo caso un profesional en el sector de dominio, calificará el resultado del desarrollo de DM, o de tipo cuantitativo, entre otras cosas, el número de detecciones de estafa o la respuesta de usuarios frente una operación publicitaria.

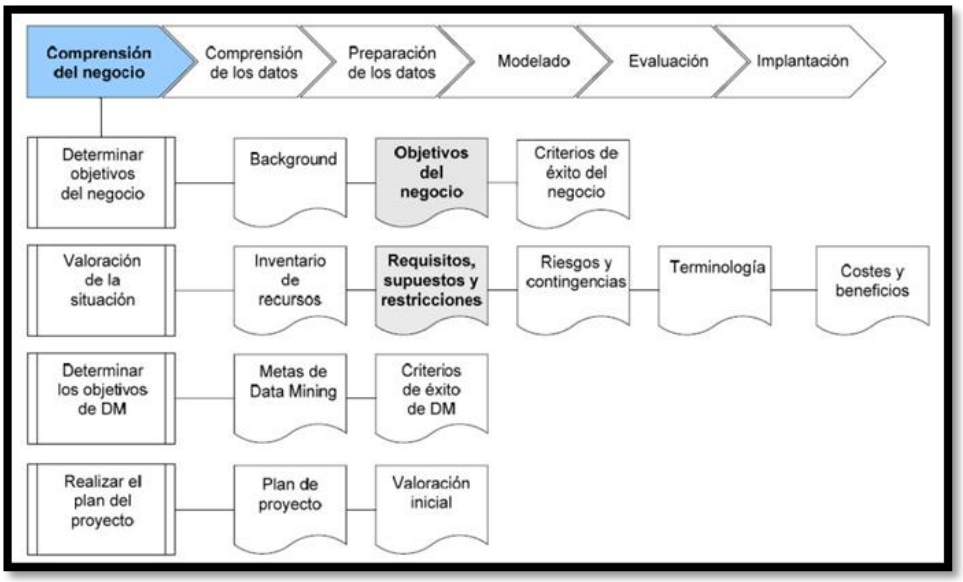


Figura 7. Fase de Comprensión del Negocio

Adaptado de “CRISP-DM” por IBM, 2012

Evaluación de la situación. En esta labor hay que calificar el estado de la circunstancia antes de comenzar el desarrollo de DM, teniendo en cuenta puntos tales como: ¿cuál es el saber previo utilizable en relación al inconveniente?, ¿se tiene la proporción de datos requerida para solucionar el inconveniente?, ¿cuál es la relación coste provecho de la aplicación de DM?, etc. En esta etapa se definen los requisitos del inconveniente, tanto en relación a negocio como en relación a Data Mining.

Determinación de los objetivos *de DM.* Esta tarea tiene como objetivo representar los objetivos del negocio en términos de las metas del proyecto de DM, como, por ejemplo, si el propósito del negocio es el avance de una operación publicitaria para aumentar la asignación de créditos hipotecarios, la misión de DM va a ser, entre otras cosas, saber el perfil de los usuarios respecto de su aptitud de endeudamiento. Producción de un plan del emprendimiento. Por último, está más reciente labor de la primera etapa de CRISP-DM, tiene como misión desarrollar un plan para el emprendimiento, que describa los procedimientos que se deben llevar a cabo y las técnicas a usar en cada paso.

A.2. Fase de Comprensión de los Datos

La segunda etapa (Figura 06), etapa de comprensión de los datos, comprende la recolección inicial de datos, con el propósito de entablar un primer contacto con el inconveniente, familiarizándose con ellos, detectar su calidad y entablar las relaciones más evidentes que permitan determinar las primeras hipotesis.

Esta etapa con las próximas dos fases, son las que denuncian el más grande esfuerzo y tiempo en un emprendimiento de DM. En la mayoría de los casos si la organización tiene una base de datos de información corporativa, es deseable hacer una exclusiva base de datos de información ad-hoc al emprendimiento de DM, ya que a lo largo del avance del emprendimiento, es viable que se generen recurrentes y abundantes accesos a la banco de información a objeto de hacer consultas y seguramente ediciones, lo cual podría crear varios inconvenientes.

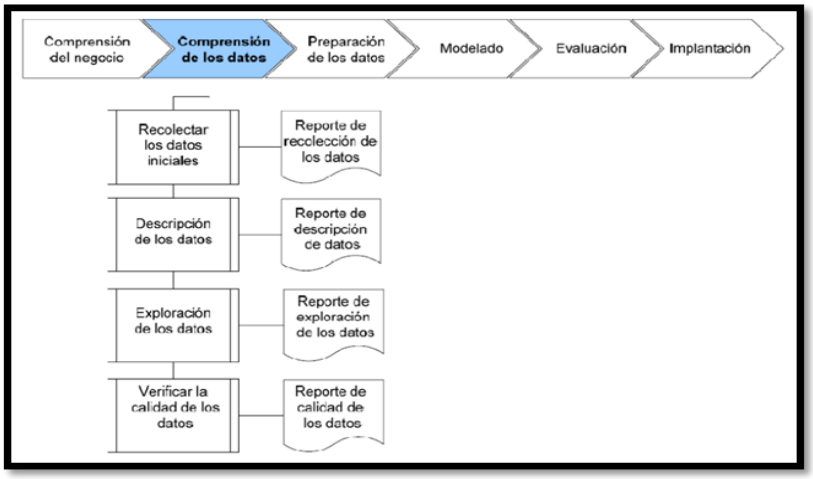


Figura 8. Fase de Comprensión de los Datos.
Adaptado de “CRISP-DM” por IBM, 2012

Las principales tareas a desarrollar en esta fase del proceso son:
Recolección de datos iniciales. La primera labor en esta segunda etapa del desarrollo de CRISP-DM, es la recolección de los datos iniciales y su adecuación para el futuro procesamiento. Esta labor tiene como propósito, llevar a cabo reportes con una lista de los datos comprados, su ubicación, las técnicas usadas en su recolección y los inconvenientes y resoluciones propios de este desarrollo.

Descripción de los datos. Luego de comprados los datos iniciales, estos tienen que ser descritos. Este desarrollo implica entablar volúmenes de datos (número de registros y campos por registro), su identificación, el concepto de cada campo y la especificación del formato inicial. Navegación de datos. Ahora, se procede a su navegación, cuyo fin es hallar una composición general para los datos. Esto implica la aplicación de pruebas estadísticas simples, que revelen características en los datos recién comprados, se crean tablas de continuidad y se edifican gráficos de organización. La salida de esta labor es un reporte de navegación de los datos.

Verificación de la calidad de los datos. En esta labor, se efectúan verificaciones sobre los datos, para saber la rigidez de los valores particulares de los campos, la cantidad y organización de los valores nulos, y para hallar valores fuera de rango, los cuales tienen la posibilidad de constituirse en ruido para el desarrollo. La iniciativa en este punto, es garantizar la completitud y rectificación de los datos.

A.3. **Fase de Preparación de los Datos**

En esta etapa y una vez efectuada la recolección inicial de datos, se procede a su elaboración para adaptarlos a las técnicas de Data Mining que se utilicen más adelante, así como técnicas de monitorización de datos, de búsqueda de relaciones entre cambiantes u otras medidas para navegación de los datos. La elaboración de datos tiene dentro las tareas en general de selección de datos a los que se va a utilizar una cierta técnica de modelado, limpieza de datos, generación de cambiantes complementarios, incorporación de diferentes orígenes de datos y cambios de formato.

Esta fase se encuentra relacionada con la fase de modelado, puesto que, en función de la técnica de modelado elegida, los datos requieren ser procesados de diferentes

formas. Es por eso las fases de elaboración y modelado interactúan de manera persistente, ilustra las superficies de que se compone ésta, e identifica sus salidas. Una especificación de las tareas involucradas en esta etapa es la siguiente: Selección de datos. En esta etapa, se selecciona un subconjunto de los datos comprados en la etapa previo, apoyándose en criterios antes establecidos en las fases anteriores: calidad de los datos en relación a completitud y rectificación de los datos y restricciones en el volumen o en los tipos de datos que están similares con las técnicas de DM seleccionadas. *Limpieza de los datos*.

Estructuración de los datos*.* Esta labor tiene dentro las operaciones de elaboración de los datos, así como la generación de nuevos atributos desde atributos ya que ya están, incorporación de nuevos registros o transformación de valores para atributos que ya están.

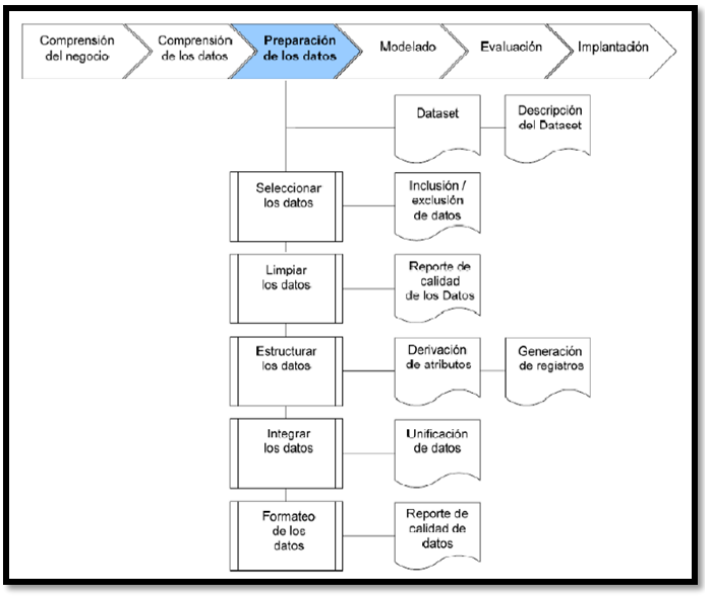


Figura 9. Fase de Preparación de Datos
Adaptado de “CRISP-DM” por IBM, 2012

Incorporación de los datos. La incorporación de los datos, implica la construcción de novedosas construcciones, desde los datos seleccionados, entre otras cosas, generación de nuevos campos desde otros que ya están, creación de nuevos registros, fusión de tablas campos o novedosas tablas donde se resumen propiedades de múltiples registros o de otros campos en novedosas tablas de resumen.

Formateo de los datos. Esta labor radica primordialmente, en la ejecución de transformaciones sintácticas de los datos sin cambiar su concepto, esto, con la iniciativa de aceptar o hacer más simple el empleo de alguna técnica de DM en especial, como entre otras cosas la reordenación de los campos y/o registros de la tabla o el ajuste de los valores de los campos a las restricciones de las utilidades de modelación (eliminar comas, tabuladores, letras y números destacables, máximos y mínimos para las cadenas de letras y números, etc.).

A.4. Fase de Modelado

En esta fase de CRISP-DM, se seleccionan las técnicas de modelado más apropiadas para el proyecto de Data Mining específico. Las técnicas a utilizar en esta fase se eligen en función de los siguientes criterios:

- Ser correcta al inconveniente.
- Disponer de datos correctos.
- Cumplir los requisitos del inconveniente.
- Tiempo correcto para conseguir un modelo.
- Conocimiento de la técnica.

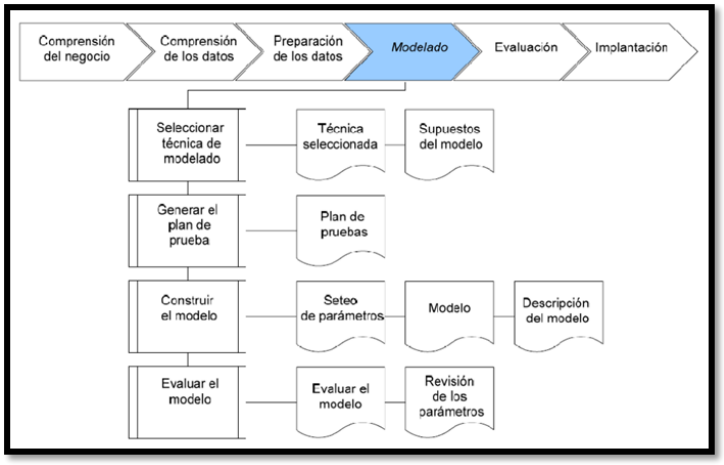


Figura 10. Fase de Modelado.
Adaptado de “CRISP-DM” por IBM, 2012

Antes al modelado de los datos, hay que saber un procedimiento de evaluación de los modelos que permita detallar el nivel de amabilidad de ellos. Luego de deducir estas tareas genéricas, se procede a la generación y evaluación del modelo. Los parámetros usados en la generación del modelo, dependen de las propiedades de los datos y de las propiedades de exactitud que se quieran conseguir con el modelo. Selección de la técnica de modelado. Esta labor radica en la selección de la técnica

de DM más correcta al tipo de inconveniente a solucionar. Para esta selección, hay que tener en cuenta el propósito primordial del emprendimiento y la relación con las utilidades de DM que ya están.

Generación del plan de prueba. Una vez constituido un modelo, hay que crear un trámite designado a evaluar la calidad y validez del mismo. Entre otras cosas, en una labor supervisada de DM como la ordenación, es habitual utilizar la causa de error como medida de la calidad. Entonces, comúnmente se separan los datos en dos conjuntos, uno de entrenamiento y otro de prueba, para después crear el modelo apoyado en el grupo de entrenamiento y medir la calidad del modelo generado con el grupo de prueba.

Construcción del Modelo. Luego de elegida la técnica, se ejecuta sobre los datos antes preparados para crear uno o más modelos. Todas las técnicas de modelado tienen un grupo de parámetros que determinan las propiedades del modelo a crear. La selección de los especiales parámetros es un desarrollo iterativo y se basa de forma exclusiva en los resultados generados. Estos tienen que ser interpretados y su desempeño justificado.

Evaluación del modelo. En esta labor, los ingenieros de DM interpretan los modelos en concordancia al conocimiento preexistente del dominio y los criterios de triunfo preestablecidos. Profesionales en el dominio del inconveniente juzgan los modelos dentro del contexto del dominio y profesionales en Data Mining aplican sus propios criterios (seguridad del grupo de prueba, perdida o ganancia de tablas, etc.).

A.5. Fase de Evaluación

En esta etapa se analiza el modelo, sabiendo el cumplimiento de los criterios de triunfo del inconveniente. Debe considerarse, además, que la fiabilidad calculada para el modelo se aplica únicamente para los datos sobre los que se llevó a cabo el análisis. Es exacto comprobar el desarrollo, sabiendo los resultados que se consiguieron, para lograr reiterar algún paso previo, en el que se haya probablemente cometido algún error.

Tener en cuenta que se tienen la posibilidad de usar múltiples utilidades para la interpretación de los resultados. Las matrices de confusión son muy empleadas en inconvenientes de categorización y consisten en una tabla que sugiere cuantas clasificaciones se hicieron para cada tipo, la diagonal de la tabla representa las clasificaciones correctas. Si el modelo generado es válido en funcionalidad de los criterios de triunfo establecidos en la etapa previo, se procede a la explotación del modelo. Las tareas involucradas en esta etapa del desarrollo son las siguientes:

Evaluación de los resultados*.* En los pasos de evaluación anteriores, se trataron causantes, así como la precisión y generalidad del modelo generado. Esta labor implica la evaluación del modelo en relación a los objetivos del negocio y busca saber si hay alguna razón de negocio para la cual, el modelo sea deficiente, o si es aconsejable evaluar el modelo, en un inconveniente real si el tiempo y limitaciones lo aceptan. Además de los resultados de manera directa relacionados con el propósito del emprendimiento, ¿es aconsejable considerar el modelo en relación a otros objetivos diferentes a los auténticos?, esto podría dar a conocer información agregada.

Proceso de revisión*.* El proceso de revisión, se refiere a calificar al proceso entero de DM, a objeto de identificar elementos que pudieran ser mejorados.

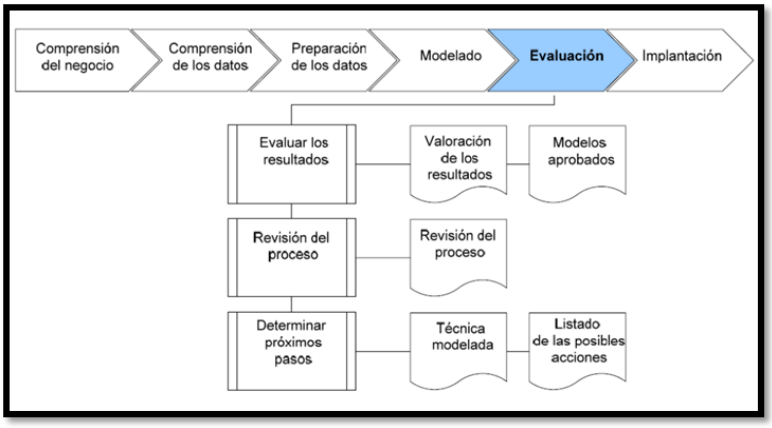


Figura 11. Fase de Evaluación
Adaptado de "CRISP-DM" por IBM, 2012

Determinación de futuras fases. Si se ha preciso que las fases hasta este instante han generado resultados satisfactorios, podría pasarse a la etapa siguiente, en caso opuesto podría decidirse por otra iteración desde la etapa de elaboración de datos o

de modelación con otros parámetros. Tiene la posibilidad de ser inclusive que en esta etapa se decida partir desde el principio con un nuevo emprendimiento de DM.

A.6. Fase de Implementación

En esta etapa, cuando el modelo fue constituido y validado, se transforma el saber obtenido en acciones dentro del desarrollo de negocio, así sea que el analista recomiende acciones fundamentadas en la observación del modelo y sus resultados, asi sea aplicando el modelo a diferentes conjuntos de datos o como parte del desarrollo, como, entre otras cosas, en análisis de peligro crediticio, descubrimiento de fraudes, etc. Por lo general un emprendimiento de Data Mining no concluye en la implementación del modelo, ya que se tienen que documentar y enseñar los resultados de forma comprensible para el usuario, con el propósito de conseguir un aumento del conocimiento. Por otro lado, en la etapa de explotación hay que garantizar el cuidado de la aplicación y la viable difusión de los resultados. Las tareas que se ejecutan en esta etapa son las siguientes:

Plan de implementación. Para llevar a cabo el resultado de DM en la organización, esta labor toma los resultados de la evaluación y concluye un plan para su utilización. Si un trámite general se ha reconocido para hacer el modelo, este trámite debe ser documentado para su posterior utilización. Vigilancia y Mantenimiento. Si los modelos resultantes del desarrollo de Data Mining son implementados en el dominio del inconveniente como parte de la rutina día tras día, es aconsejable elaborar tácticas de vigilancia y cuidado para ser aplicadas sobre los modelos. El reabastecimiento generado por la vigilancia y cuidado tienen la posibilidad de señalar si el modelo es usado apropiadamente.

Informe Final. Es la conclusión del emprendimiento de DM llevado a cabo. En relación del plan de utilización, este reporte puede ser sólo un comprendido de los puntos indispensables del emprendimiento y la vivencia lograda o puede ser una exhibición final que integre y explique los resultados logrados con el emprendimiento. Revisión del proyecto: En este punto se analiza qué fue lo preciso y qué lo incorrecto, qué es lo que se realizó bien y qué es lo que se necesita hacer mejor.

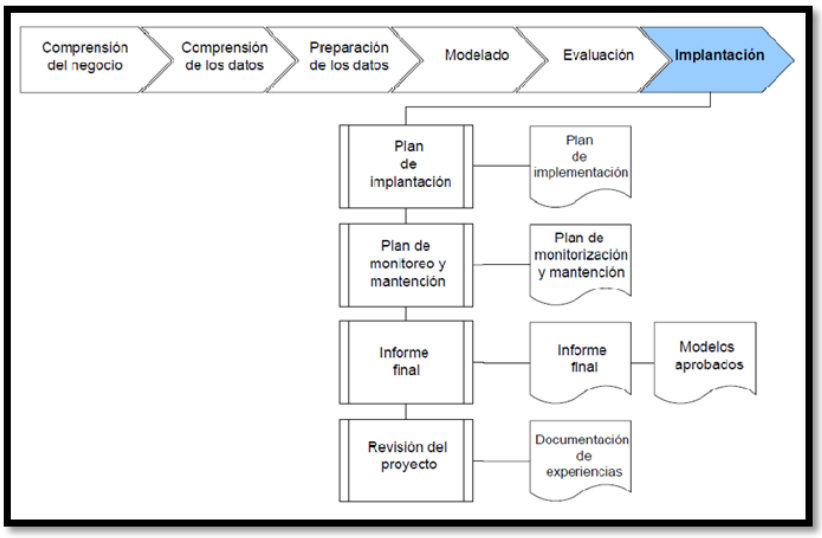


Figura 12. Fase de Implementación
Adaptado de "CRISP-DM" por IBM, 2012

2.2.5. Toma de Decisiones Académicas

La toma de decisiones es un desarrollo por medio del cual se ejecuta una selección entre distintas opciones o maneras de solucionar o combatir una cierta circunstancia que se puede enseñar, bien sea a nivel personal, laboral u organizacional. Y radica principalmente en elegir una alternativa, que es considerada la más correcta, entre una sucesión de configuraciones accesibles, con el objetivo de solucionar un preciso inconveniente que se tiene hoy en día, o para combatir un potencial inconveniente que se puede enseñar en un preciso instante.

En la situación de las organizaciones, la toma de elecciones es una herramienta imprescindible para su recurrente desarrollo, y en ese sentido la administración debe usar la racionalidad para tomar selecciones acertadas y elegir el más considerable sendero a seguir según las diferentes selecciones para batallar una alguna situación y sostener la operatividad y la efectividad de la organización, aun en oportunidades de indecisión. Y en la circunstancia particular de una institución didáctica, la toma de selecciones es de mucha consideración para desarrollar las funciones administrativas de idealización, de organización, y de control de las diferentes ocupaciones que se llevan a cabo en una institución didáctica para el recurrente avance de los procesos académicos y administrativos de la institución.

Proceso de Toma de Decisiones

La necesidad de tomar decisiones rápidamente en un mundo cada vez más complejo y en continua transformación, puede llegar a ser muy desconcertante, por la imposibilidad de asimilar toda la información necesaria para adoptar la decisión más adecuada. Todo ello nos conduce a pensar que el tomar decisiones supone un proceso mental, que lleva en sí mismo los siguientes pasos

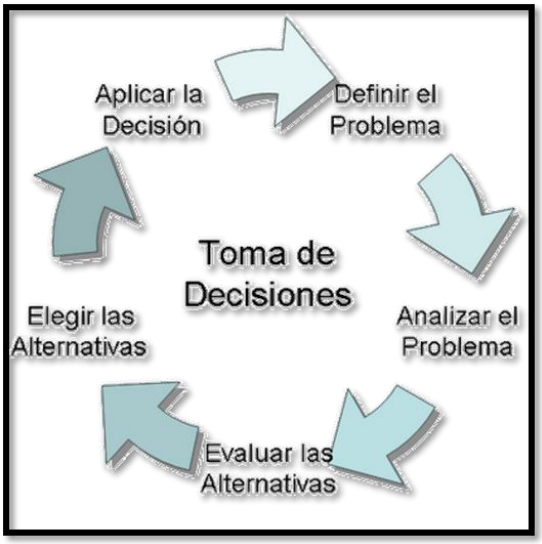


Figura 13. Proceso de Toma de Decisiones
Adaptado de "CRISP-DM" por IBM, 2012

- **Identificación del problema**: debemos admitir cuando nos encontramos frente un inconveniente para buscar elecciones al mismo. En este primer escalón debemos cuestionarnos, ¿qué hay que elegir?

- **Análisis del problema:** en este paso hemos de saber las causas del inconveniente y sus secuelas y agarrar la máxima información viable sobre el mismo. En esta oportunidad la cuestión a solucionar es, ¿Cuáles son las configuraciones probables?
- **Evaluación o estudio de opciones o alternativas:** aquí nos debemos centrar en detectar las probables resoluciones al inconveniente o tema, de esta forma como sus probables secuelas. Nos debemos preguntar, ¿cuáles son las virtudes e problemas de cada opción?
- **Selección de la mejor opción:** una vez analizadas todas las configuraciones o elecciones probables, debemos elegir la que no se ve más favorable y correcta. Observamos como aquí está implicada en sí misma una elección, en esta oportunidad nos preguntamos ¿cuál es la preferible alternativa?

- **Poner en práctica las medidas tomadas:** una vez tomada la elección debemos llevarla a la costumbre y ver su evolución. Aquí reflexionamos sobre ¿es precisa la elección?
- **Finalmente evaluamos el resultado:** en esta más reciente etapa debemos tener en cuenta si el inconveniente se ha resuelto acorde a lo previsto, analizando los resultados para cambiar o volver a pensar el desarrollo en los puntos necesarios para hallar el propósito pretendido. En esta etapa nos preguntamos, ¿la elección tomada produce los resultados deseados?

La toma de elecciones académicas constituye hoy en día, en el contexto universitario, un desarrollo complejo no sólo por la enorme diversidad de configuraciones que ya están, sino además por los resultados que se derivan de la selección realizada. Hay que investigar la diversidad de cambiantes personales, académica y contextual que intervienen e influyen en la selección de la toma de elección. Desde los criterios que se tienen presente, el trámite que se sigue o el nivel de agrado con la elección tomada, se transporta a cabo un análisis detallado. Los resultados no sólo han de ser útil para comprender mejor cómo afronta el estudiante este suceso, sino además para ingresar las medidas oportunas de carácter orientador que posibiliten esa selección.

Minería de Datos en la Educación

La minería de datos en la educación es una disciplina emergente que desarrolla métodos para explorar los datos que vienen de entornos educativos, y los usa para entender mejor a los estudiantes y los entornos en los que aprenden

Ventajas respecto a los Paradigmas tradicionales de Investigación educacional

La minería de datos educacionales da varias virtudes comparándola con los paradigmas más habituales de exploración relativa a la educación, como experimentos de laboratorio, estudios sociológicos o exploración de interfaz. En especial, la construcción de repositorios públicos de datos educacionales creó una base que hace viable la minería de datos educacionales. En especial, los datos de estos repositorios son completamente válidos (ya que son datos reales sobre el desempeño y estudio de alumnos reales, en ambientes educacionales, tomados en tareas de aprendizaje), y cada vez más de forma sencilla accesibles para empezar una exploración.

Principales enfoques

Hay una extensa variedad de procedimientos empleados comúnmente en el tema de la educación en la minería de datos. Estos procedimientos están comprendidos en las siguientes categorías: predicción, agrupamiento, minería de relaciones, inferencia por medio de modelos, y destilación de datos para la interpretación de parte de un ser humano. Las tres primeras categorías son universales para diferentes tipos de minería de datos (aunque en algunas ocasiones con diferentes nombres). Las categorías cuarta y quinta alcanzan una especial consideración dentro de la minería de datos educacionales.

Predicción

En predicción, el propósito es desarrollar un modelo que logre inferir una variable desde alguna conjunción de otras cambiantes dentro en los datos. La predicción necesita etiquetas para la variable de salida para un grupo de datos con limite, donde una etiqueta suponga una información confiable sobre el valor de la variable de salida en casos particulares.

Agrupamiento

En agrupamiento, el objetivo es encontrar puntos de datos que se agrupen de manera natural, repartiendo el conjunto original de datos en un conjunto de ‘clusters’. El agrupamiento es particularmente útil en casos donde las categorías más comunes de los datos no son conocidas. Si un conjunto de clústeres es óptimo, en cada categoría, cada punto será más similar a los puntos pertenecientes a su clúster que a puntos pertenecientes a otros grupos. Los clústeres pueden ser creados con distinta granularidad

Minería de Relaciones

En la minería de relaciones, el propósito es conocer relaciones entre cambiantes en un grupo de datos con un enorme conjunto de cambiantes. Una manera de hacer esto es localizando las cambiantes más poderosamente similares con una exclusiva variable de interés, o además por medio del hallazgo de las relaciones más fuertes entre dos cambiantes. De forma general, hay 4 tipos de minería de relaciones: minería de reglas de organización, minería de correlación, minería de patrones de secuencias y minería de datos causales.

Descubrimiento mediante Modelos

En el descubrimiento a través de modelos, se transporta a cabo un modelo a través de predicción, agrupamiento o, en varias oportunidades, ingeniería del conocimiento (usando métodos de razonamiento humano en vez de métodos automatizados). Este modelo es entonces utilizado como un ingrediente en otro análisis, como predicción o minería de datos. En la circunstancia de predicción, las conjeturas hechas por el modelo creado son usadas como cambiantes de entrada en la predicción de una única variable.

Destilado de datos

Otro sector de interés en la minería de datos educacionales es la destilación de datos para la interpretación humana. En algunas ocasiones, los humanos tienen la posibilidad de hacer inferencias sobre datos cuando éstos son presentados como corresponde, que están más allá del punto de mira inmediato de los procedimientos de minería de datos completamente automatizados. Los procedimientos en este sector de minería de datos educacionales son de información y procedimientos de monitorización. De todos métodos, las visualizaciones más usadas en el campo educacional tienden a ser diferentes a las usadas para la resolución de inconvenientes de monitorización de la información, debiéndose a la composición específica y el concepto embebido en esa composición, comúnmente que se encuentra en datos educacionales.

Aplicaciones de la minería de datos en educación

- Mejorar los modelos del estudiante. La minería de datos se ha usado, por ejemplo, para determinar si un estudiante truca el sistema o si está aburrido o frustrado en un curso online. Una versión más amplia de los modelos del estudiante va más allá del software educativo y se propone determinar, por ejemplo, los factores de los que depende si alguien va a abandonar sus estudios.
- Descubrir y mejorar la estructura del conocimiento en un dominio.
- Estudiar el apoyo pedagógico, por ejemplo, mediante la descomposición del aprendizaje.

CAPÍTULO III

CONSTRUCCIÓN DE LA SOLUCIÓN

3.1. ESTUDIO DE FACTIBILIDAD

3.1.1. Factibilidad Técnica

La Factibilidad Técnica consistió en hacer una evaluación de la tecnología que existe en la organización, este estudio estuvo designado a juntar información acerca de los elementos técnicos que tiene la organización y la oportunidad de utilizar los mismos en el avance y utilización del sistema propuesto y de ser primordial, los requerimientos tecnológicos que tienen que ser comprados para el avance y puesta en marcha del sistema.

En tal sentido, la tecnología requerida para la habilitación y construcción de la aplicación es:

- Software de minería de Datos WEKA.
- Software Modelador de Procesos: BIZAGI
- Paquete de Ofimática: Office 2010

3.1.2. Factibilidad Operativa

La Factibilidad Operativa facilita adivinar, si se va a poner en marcha el sistema propuesto, utilizando los beneficios que da, a todos los individuos comprometidos con el mismo, ya sean los que interactúan en forma directa con este, como además esos que reciben información producida por el sistema. Por lo cual el sistema es posible operativamente por las siguientes causas:

- El correcto funcionamiento del sistema en cuestión, siempre estará supeditado a la capacidad de los secretarios académicos encargados de dicha tarea.
- La necesidad y deseo de un cambio en el sistema actual, expresada por los secretarios académicos, llevó a la aceptación para implementar un nuevo sistema, que, de una manera más sencilla y amigable, cubra todos sus requerimientos, expectativas y proporciona la información en forma oportuna y confiable.
- Basándose en las entrevistas y conversaciones sostenidas con los secretarios académicos se demostró que estos no representan ninguna oposición al cambio, por lo que el sistema es factible operacionalmente.

3.1.3. Factibilidad Económica

A continuación, se muestra un estudio que dio como resultado la factibilidad económica del avance del nuevo servicio de información. Se concluyeron los elementos para desarrollar, establecer, y sostener en operación el sistema planificado, realizando una evaluación donde se puso de manifiesto la estabilidad que existe entre los costos intrínsecos del sistema y los provecho que se derivaron de éste, lo cual permitió ver de una forma más precisa las bondades del sistema propuesto.

A continuación, presentamos los recursos necesarios para la implementación del proyecto.

Tabla 13
Factibilidad Económica

RUBROS	UNIDAD DE MEDIDA	PRECIO UNITARIO	CANTIDAD	TOTAL
1. Recursos humanos				
Ofelia Rodriguez Miranda	Persona	2500	1	2500
2. Recursos Materiales				
Papel	Millar	24	2	48
Lapicero	Global	1.5	30	45
Engrapador	Global	8	1	8
Folder	Global	0.5	10	5
Fotocopias	Global	0.1	200	20
3. Hardware				
Computadora	Global	2500	1	2500
Impresora	Global	500	1	500
4. Software				
Microsoft Office 2010	Global	500	1	500
5. otros Gastos	Global	300	1	300
Total Presupuesto				**6426**

3.2. FASE DE COMPRENSIÓN DEL NEGOCIO

3.2.1. Determinar objetivos del negocio

1.1 Visión, Misión y Principios

Teniendo en cuenta la experiencia ganada en los cinco primeros años de gestión institucional supervisada y acreditada por CONAFU, hemos trazado el guión institucional que nos oriente a mayores y mejores éxitos en los primeros nueve años de gestión autónoma, aunque también acreditada por los entes reguladores de la calidad tanto nacional como global.

Visión

La Universidad Autónoma del Perú será reconocida por su excelencia en la formación de profesionales y académicos en base al compromiso de los miembros de su comunidad.

Misión

Formamos personas, profesionales y académicos competitivos para generar valor a las organizaciones de la sociedad, a la ciencia y al desarrollo sustentable, con un capital humano líder y comprometido.

Principios

1) Búsqueda de la realidad y la excelencia, con fortalecimiento de valores y respeto al avance sostenible.
2) Pluralismo e independencia de pensamiento, de crítica, de expresión y de cátedra, con lealtad a la Constitución y Legislación Peruana como al propio Estatuto.
3) Creatividad e creatividad para entrenar una importante cultura de evaluación, acreditación, originalidad y generación de elementos propios.
4) Autenticidad, afirmando siempre que solo somos y vamos a ser lo que nuestra Red social y la Sociedad aprecien y valoren que hacemos bien.
5) Competitividad de nuestro capital humano para interpretar en qué debemos hacer mejor y cómo llevarlo a cabo eficiente y eficientemente para estar siempre entre los especiales.

3.2.2. Objetivos del Negocio

1. Continuar la acreditación institucional e comenzar la de carreras expertos.
2. Mejorar el avance institucional subiendo la oferta de formación profesional e iniciando la formación técnica y de postgrado.
3. Elevar la solidez institucional basado en hacer mejor la retención académica y su puntualidad de pago de pensiones educativas.
4. Desarrollar el capital humano para la optimización persistente de nuestra competitividad y diferenciación institucional.
5. Elevar la oferta de formación continua para hacer mejor la generación de capital y de utilidades.

3.2.3. Resultados esperados

A partir del segundo semestre académico del año 2020 debemos tener como logro sostenido lo siguiente:

- Brindar siempre increíble atención a todos nuestros alumnos y docentes para garantizar el logro de las misiones de avance progresivo anual.
- Mantener una relación estudiante/docente entre 28 y 32. Tener el 45% de docentes con grados académicos avanzados.
- Tener el 10% de docentes formalmente en carrera regular. Los docentes tienen que hacer siempre exploración científica y enseñar dos trabajos concluidos y publicados por carrera todos los años, desde el egreso de su primera promoción.
- Todos los egresados se titularán sustentando y aprobando una proposición que va a ser el resultado de un trabajo de exploración, las mismas que tienen que ofrecer origen a artículos investigadores popularizados oportunamente.

a) La proyección popular de cada carrera debe ser persistente y ofrecer consideración al compromiso popular.
b) La herramienta por generación de elementos propios, diferentes de las tasas educativas, debe aceptar atender más del 15% de los costos corrientes.
c) Mejora persistente en acreditación Institucional y de Carreras Expertos al menos en todo el país.
d) Permanente seguridad económica y financiera con increíble Administración Académica, Administrativa y Económica. Increíble atención a trabajadores y grupos de interés.

3.2.4. Factores críticos de éxito del negocio

- Acreditación Universitaria
- Deserción en un rango menor del 10%
- Educación de calidad

3.2.5. Inventario de recursos.

- Computadora para desarrollo y pruebas.
- Impresora.
- Software Weka.
- Suite Office 2010.

3.2.6. Fuentes de información

- Sistema de Información Trilce.
- Manuales de Weka.
- Documentación de investigaciones de Data Mining.
- Algoritmos de Clustering.
- La principal fuente de datos para llevar a cabo esta investigación la constituyeron los estudiantes del I ciclo de la carrea de Ingeniería de Sistemas.

3.2.7. Recursos humanos

Auspiciador: jefe de la carrera de Ing. De sistemas

Usuarios Involucrados:

- Jefa del área de calidad educativa de la universidad
- Directora de formación general
- Docentes de los cursos.

3.2.8. Requerimientos, expectativas y restricciones

Existe restricciones con respecto a la información obtenida de los estudiantes sobre sus notas, para eso se está coordinando con el Jefe de carrera para que proporcione la información.

LA información proporcionada puede tener errores al momento de ser adquirida, por lo tanto se tiene que validar en el momento de su recojo.

Los docentes pueden malinterpretar los resultados obtenidos.

3.2.9. Riesgos y contingencias

- Encuestas mal desarrolladas
- El auspiciador no dispone de tiempo para las reuniones.
- Desconocimiento del uso de la herramienta WEKA.
- El jefe de carrera no autoriza la realización del levantamiento de información.

3.2.10. Costos y beneficios

Costo de Recolección de Datos: 200 Soles.
Costo del desarrollo y análisis: 2500 soles.

3.2.11. Determinar los objetivos de DM

El principal objetivo del DM es determinar las características del perfil de aprendizaje de los estudiantes del primer ciclo para, en función de él, adecuar las estrategias de enseñanza de los docentes de la carrera.

3.3. FASE DE COMPRENSIÓN Y PREPARACIÓN DE LOS DATOS

Para la recolección de los datos se aplicó el test e Felder que se expone en los anexos. Después de aplicado el Test se procedió a preprocesamiento para la limpieza de los datos por medio del reemplazo de valores faltantes por el signo “?”, además se removieron algunas inconsistencias por medio de correcciones manuales, así como la presencia de espaciados insignificantes y otros fallos de tecleo.

Ya que se trabajó con un exclusivo archivo no fue primordial hacer tareas de incorporación ni de transformación. Se llevó a cabo una exclusiva labor de reducción de datos que consistió en remover los atributos: nombre del alumno, por ser sin importancia para los objetivos de este desarrollo de hallazgo de conocimiento, y las respuestas a los ítems 8 y 16, por corresponder a la extensión procesamiento.

Se generaron variables como NOMBRE Y APELLIDO, EDAD CICLO, TURNO, SECCION, SEXO, TIPO, COLEGIO, DISTRITO, PREG 1, PREG 2, PREG 3, PREG 4, PREG 5, PREG 6, PREG 7, PREG 8, PREG 9, PREG 10, PREG 11, PREG 12, PREG 13, PREG 14, PREG 15 PREG 16, PREG 17, PREG 18, PREG 19, PREG 20, PREG 21, PREG 22, PREG 23, PREG 24, PREG 25, PREG 26, PREG 27, PREG 28, PREG 29, PREG 30, PREG 31, PREG 32, PREG 33, PREG 34, PREG 35, PREG 36, PREG 37, PREG 38, PREG 39, PREG 40, PREG 41, PREG 42, PREG 43, PREG 44

3.4. FASE DE MODELADO

3.4.1. Seleccionar técnica de modelado

Clustering con K-Medias

3.4.2. Construir el modelo

Con estos datos, se procedió a cargar la data de la encuesta en el programa WEKA, para lograr llevar a cabo tareas de minería de datos, que tienen dentro las de tipo descriptivo, como las que soportan el presente trabajo. En esta situación especial se aplicó análisis de clúster para detectar los subgrupos homogéneos de La gente encuestada. Como puede apreciarse en los resultados (Figura 4), para descubrir los estilos de estudio dominantes se definió en dos el número de clúster a crear.

=== Run information ===

Scheme: weka.clusterers.SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Relation: test1

Instances: 42

Attributes: 44

PREG 1

PREG 2

PREG 3

PREG 4

PREG 5

PREG 6

PREG 7

PREG 8

PREG 9

PREG 10

PREG 11

PREG 12

PREG 13

PREG 14

PREG 15

PREG 16

PREG 17

PREG 18

PREG 19

PREG 20

PREG 21

PREG 22

PREG 23

PREG 24

PREG 25

PREG 26

PREG 27

PREG 28

PREG 29

PREG 30

PREG 31

PREG 32

PREG 33

PREG 34

PREG 35

PREG 36

PREG 37

PREG 38

PREG 39

PREG 40

PREG 41

PREG 42

PREG 43

PREG 44

Test mode: evaluate on training data

=== Clustering model (full training set) ===

kMeans
======

Number of iterations: 6
Within cluster sum of squared errors: 577.0
Missing values globally replaced with mean/mode

Cluster centroids:

Attribute	Full Data (42)	Cluster# 0 (33)	1 (9)
PREG 1	A	A	A
PREG 2	A	A	B
PREG 3	A	A	B
PREG 4	B	B	B
PREG 5	A	A	B
PREG 6	A	A	B
PREG 7	A	A	A
PREG 8	A	A	B
PREG 9	A	A	A
PREG 10	A	A	A
PREG 11	A	A	A
PREG 12	A	A	B
PREG 13	A	A	B
PREG 14	B	B	B
PREG 15	B	B	A
PREG 16	A	A	A
PREG 17	B	B	B
PREG 18	A	A	B
PREG 19	A	A	A

PREG 20	A	A	A
PREG 21	B	B	A
PREG 22	B	B	B
PREG 23	A	A	A
PREG 24	A	A	B
PREG 25	B	B	B
PREG 26	B	B	A
PREG 27	A	A	A
PREG 28	B	B	A
PREG 29	A	A	A
PREG 30	A	A	B
PREG 31	B	A	B
PREG 32	A	A	A
PREG 33	A	A	B
PREG 34	B	B	B
PREG 35	A	A	A
PREG 36	A	A	B
PREG 37	A	A	B
PREG 38	A	A	A
PREG 39	A	A	A
PREG 40	B	A	B
PREG 41	A	A	A
PREG 42	A	A	B
PREG 43	A	A	A
PREG 44	A	A	B

Time taken to build model (full training data) : 0.02 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 33 (79%)
1 9 (21%)
2

3.5. FASE DE EVALUACIÓN

3.5.1. Evaluar los resultados

Un estilo de aprendizaje se define como unas preferencias o resistencias características en el modo en que el alumno adquiere y procesa la información (Felder, 1988) y ello determina la única forma de aprender que tiene cada alumno. En el modelo de Felder (Felder, 1988), se definen diferentes dimensiones relativas a cómo el alumnado procesa la información, de modo que cada dimensión tiene dos posibles valores.

Tabla 14

Comportamiento de los estudiantes de acuerdo con su aprendizaje

Dimensión	Estilos	Descripción
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
	Reflexivo	Prefiere pensar las cosas a fondo antes de actuar.
Percepción	Sensorial	Se orienta hacia hechos Percepción y datos concretos.
	Intuitiva	Prefiere lo conceptual y teórico a lo concreto.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
	Verbal	Tiene mayor éxito cuando oye o lee la información con palabras.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada.
	Global	Prefiere ver la imagen de conjunto primero.

A partir de las características definidas por Felder y Silverman, es posible determinar las correspondencias entre los estilos de aprendizaje y las posibles respuestas del test; las cuales se ilustran en la tabla 14

Tabla 15

Correspondencia entre los estilos de aprendizaje respuestas del test de Felder

Estilos de Aprendizaje	Código Respuestas	Consideradas
Activo	AC	P1=a, P5=a, P9=a, P13=a, P17=a, P21=a, P25=a, P29=a, P33=a, P37=a, P41=a.
Reflexivo	R	P1=b, P5=b, P9=b, P13=b, P17=b, P21=b, P25=b, P29=b, P33=b, P37=b, P41=b.
Sensorial	SN	P2=a, P6=a, P10=a, P14=a, P18=a, P22=a, P26=a, P30=a, P34=a, P38=a, P42=a.
Intuitivo	I	P2=b, P6=b, P10=b, P14=b, P18=b, P22=b, P26=b, P30=b, P34=b, P38=b, P42=b.
Visual	V	P3=a, P7=a, P11=a, P15=a, P19=a, P23=a, P27=a, P31=a, P35=a, P39=a, P43=a.
Verbal	VB	P3=b, P7=b, P11=b, P15=b, P19=b, P23=b, P27=b, P31=b, P35=b, P39=b, P43=b.
Secuencial	SC	P4=a, P8=a, P12=a, P16=a, P20=a, P24=a, P28=a, P32=a, P36=a, P40=a, P44=a.
Global	G	P4=b, P8=b, P12=b, P16=b, P20=b, P24=b, P28=b, P32=b, P36=b, P40=b, P44=b.

A partir de este conocimiento, se procedió a reemplazar cada uno de los atributos incluidos en los dos clústeres por el código del estilo de aprendizaje con que se vincula. Una vez que se hizo la sustitución, los clústeres quedaron como se muestra en la tabla 15. Se definió la nomenclatura para cada estilo como sigue: AC=activo, R=reflexivo; SN=sensorial¸ I=intuitivo; V=visual, VB=verbal; SC=secuencial, G=global.

Tabla 16

Representación de los atributos por los estilos correspondientes

PORCENTAJE	CLUSTER	PREG1	PREG2	PREG3	PREG4	PREG5	PREG6	PREG7	PREG8	PREG9	PREG10	PREG11	PREG12	PREG13	PREG14	PREG15	PREG16	PREG17	PREG18	PREG19	PREG20	PREG21	PREG22	PREG23	PREG24	PREG25	PREG26	PREG27	PREG28	PREG29	PREG30	PREG31	PREG32	PREG33	PREG34	PREG35	PREG36	PREG37	PREG38	PREG39	PREG40	PREG41	PREG42	PREG43	PREG44
79%		AC	SN	V	G	AC	SN	V	SC	AC	SN	V	SC	AC	I	VB	SC	R	SN	V	SC	R	I	V	SC	R	I	V	G	AC	SN	V	SC	AC	I	V	SC	AC	SN	V	SC	AC	SN	V	SC
	Cluster 0	A	A	A	B	A	A	A	A	A	A	A	A	A	B	B	A	B	A	A	A	B	B	A	A	B	B	A	B	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A
21%		AC	I	VB	G	R	I	V	G	AC	SN	V	G	R	I	V	SC	R	I	V	SC	AC	I	V	G	r	SN	V	SC	AC	I	VB	SC	R	I	V	G	R	SN	V	G	AC	I	V	G
	Cluster 1	A	B	B	B	B	B	A	B	A	A	A	B	B	B	A	A	B	B	A	A	A	B	A	B	B	A	A	A	A	B	B	B	B	B	A	B	B	A	A	B	A	B	A	B

A fin de dilucidar la combinación de estilos de aprendizaje representada por cada clúster, se procedió a contar la cantidad de ocurrencia de cada estilo, con lo cual resultan las combinaciones que muestra la tabla 16

Tabla 17
Combinación de estilos de aprendizaje representada por cada clúster

DIMENSION	PROCESAMIENTO		PERCEPCION		ENTRADA		COMPRENSION	
CLUSTER	AC	R	SN	I	V	VB	SC	G
CLUSTER0	8	3	7	4	10	1	9	2
CLUSTER1	5	6	3	8	9	2	4	7

Analizando la tabla anteriormente mostrada, pueden definirse los patrones de estilos de aprendizaje descubiertos combinando 16l Cuestionario sobre el Índice de Estilos de Aprendizaje de Felder y la minería de datos. De este modo, puede inferirse que el primer clúster (0), que representa al 79% de la población, tiene una combinación predominante de estilo de aprendizaje que corresponde con la combinación activo-sensitivo-visual-global. El segundo clúster (1) por su parte, agrupa una combinación reflexivo-intuitivo-visual-global, en un 21%.

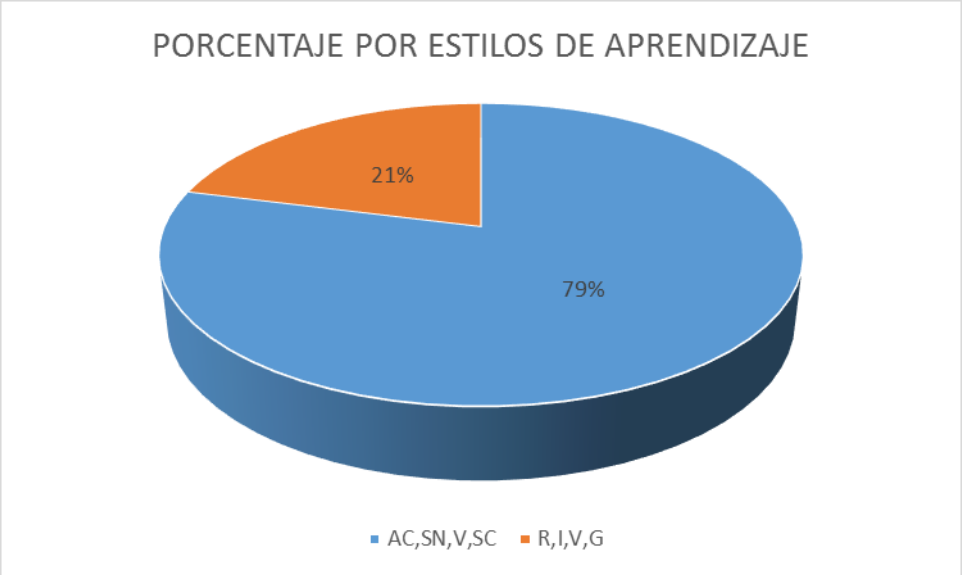


Figura 14. Combinación de estilos de aprendizaje por clúster

3.6. FASE DE IMPLANTACIÓN

3.6.1. Plan de implantación

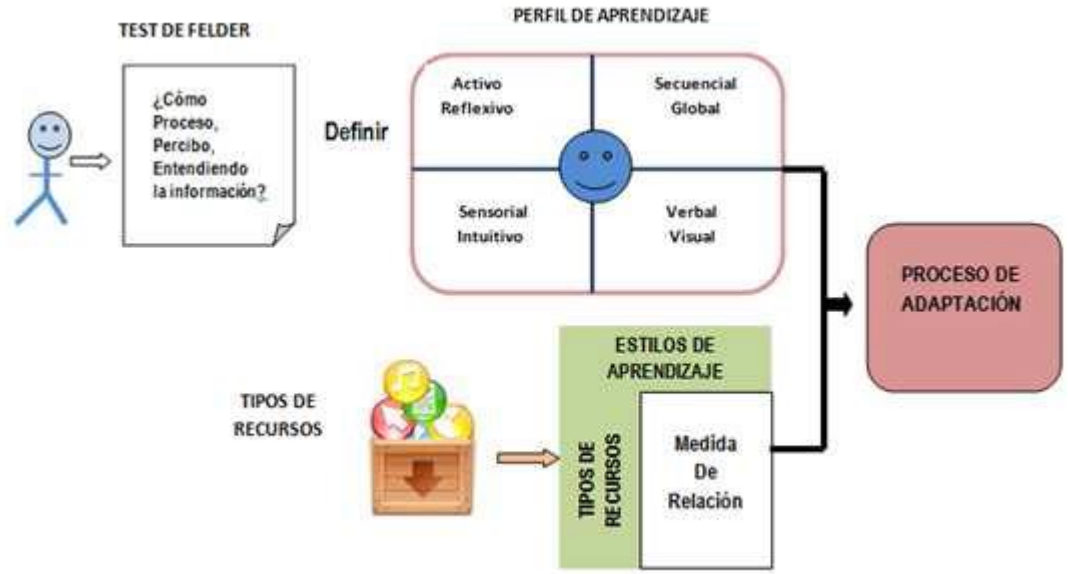
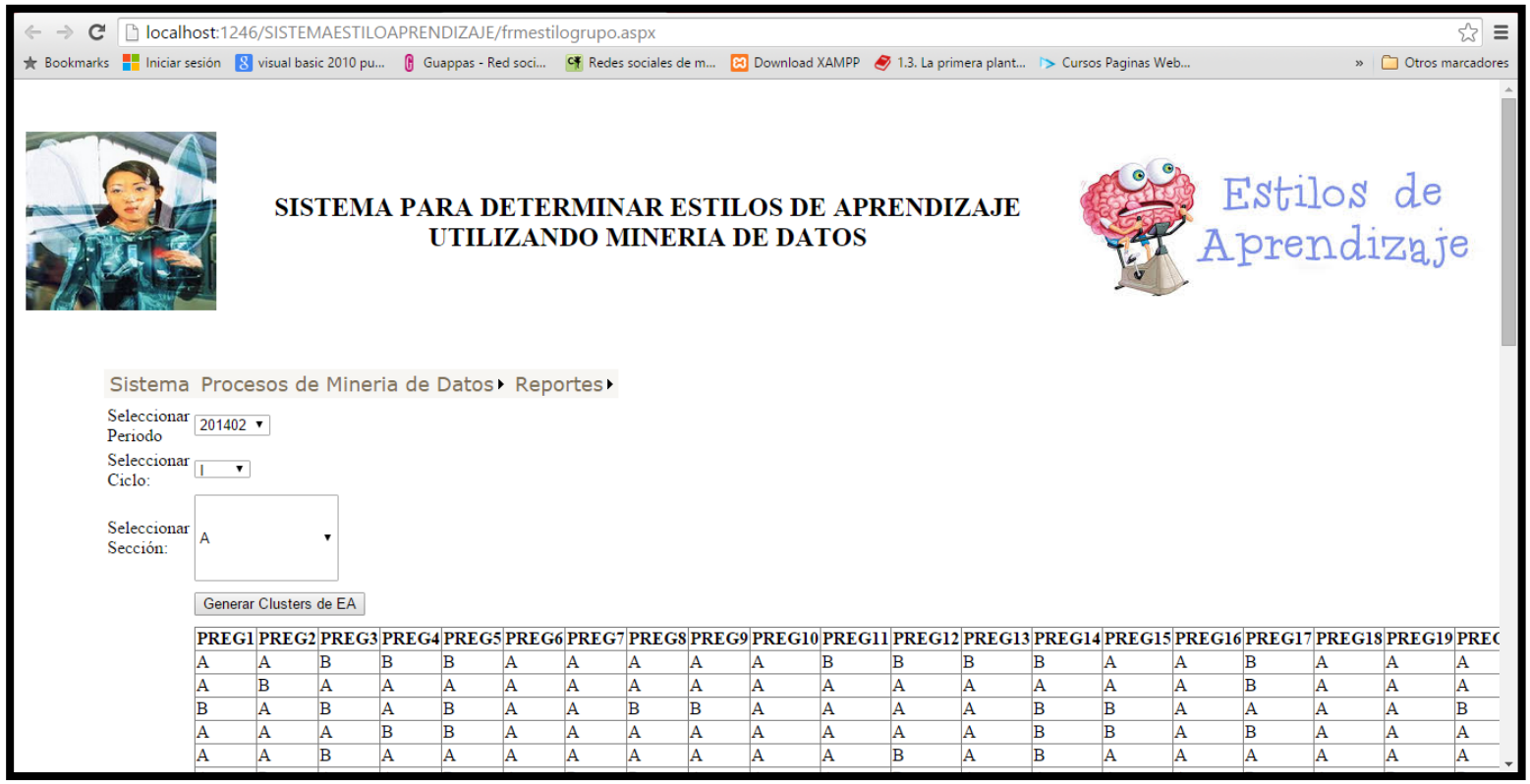


Figura 15. Esquematización preliminar de las tres primeras fases del proceso de adaptación

Para la implantación del modelo se desarrollará una aplicación web que permita automatiza el proceso anterior.

- Página Web para generar los clústeres:



PREG1	PREG2	PREG3	PREG4	PREG5	PREG6	PREG7	PREG8	PREG9	PREG10	PREG11	PREG12	PREG13	PREG14	PREG15	PREG16	PREG17	PREG18	PREG19	PREC
A	A	B	B	B	A	A	A	A	A	B	B	B	B	A	A	B	A	A	A
A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A
B	A	B	A	B	A	A	B	B	A	A	A	A	B	B	A	A	A	A	B
A	A	A	B	B	A	A	A	A	A	A	A	A	B	B	A	B	A	A	A
A	A	B	A	A	A	A	A	A	A	A	B	A	B	A	A	A	A	A	A

Figura 16. Página Web para generar estilos de aprendizaje por sección

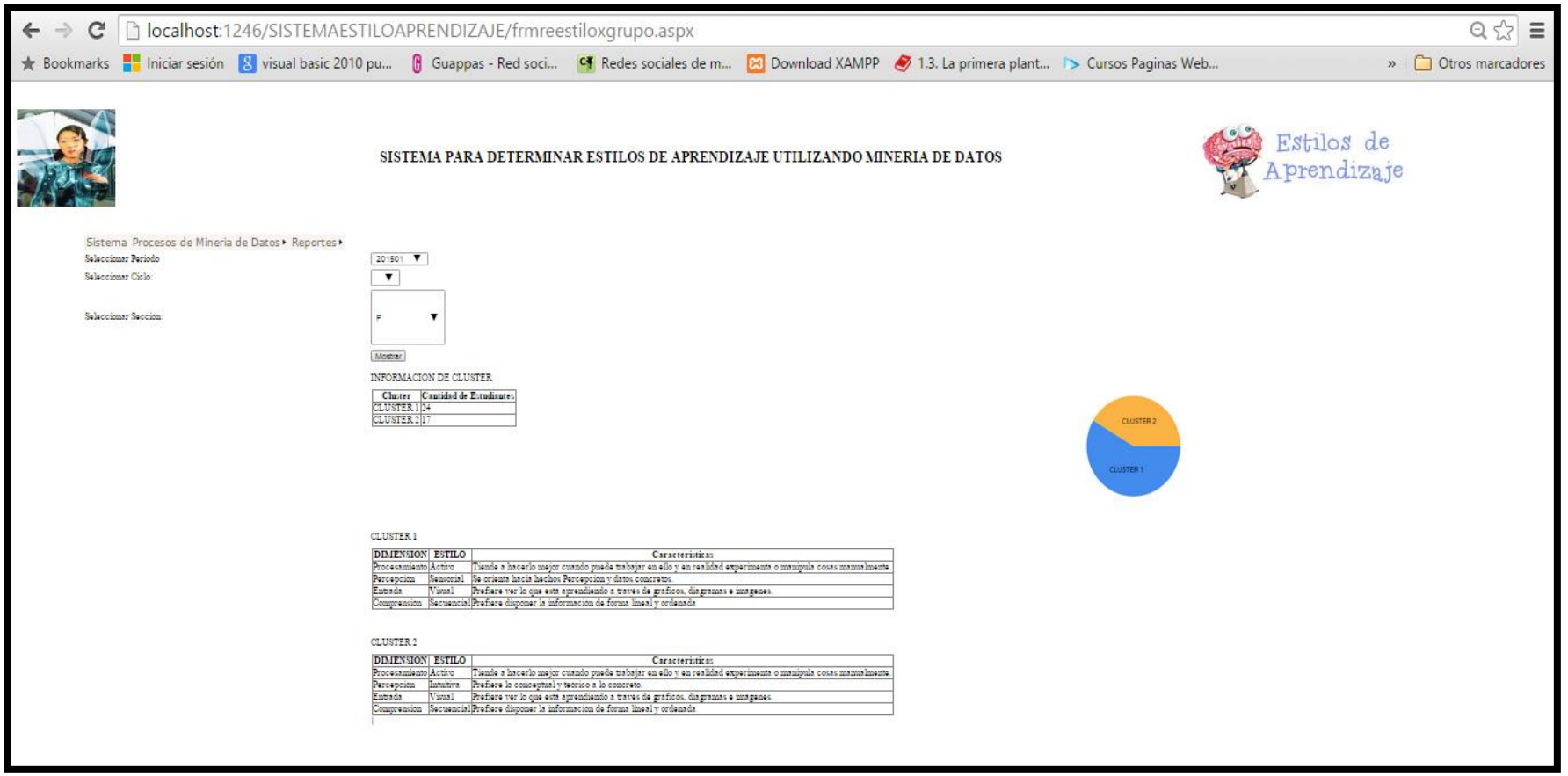


Figura 17. Página Web para mostrar los estilos de aprendizaje por sección

CAPÍTULO IV

ANÁLISIS DE RESULTADOS Y CONTRASTACIÓN DE LA HIPÓTESIS

4.1. POBLACIÓN Y MUESTRA

4.1.1 Población

N=Secciones de la Carrera de Ing. De sistemas

4.1.2 Muestra

n=30 secciones del de la carrera de Ingeniería de sistemas.

4.2. NIVEL DE CONFIANZA Y GRADO DE SIGNIFANCIA

100% ya que se toma todos los estudiantes del primer ciclo.

4.3. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

4.3.1. Resultados Genéricos

4.3.2. Resultados de Clústeres

SECCCION A

Tabla 18
Cantidad de Estudiantes por Clúster de la Sección A

Cluster	Cantidad de Estudiantes
CLUSTER 1	19
CLUSTER 2	26

Tabla 19
Estilos de Aprendizaje del Clúster 1

CLUSTER 1		
DIMENSION	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Sensorial	Se orienta hacia hechos Percepción y datos concretos.
Entrada	Visual	Elige ver lo que está aprendiendo por medio de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

Tabla 20
Estilos de aprendizaje del Clúster 2

CLUSTER 2		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Intuitiva	Elige lo conceptual y teórico a lo preciso.
Entrada	Visual	Elige ver lo que está aprendiendo por medio de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Elige contar con la información de manera lineal y organizada

SECCIÓN B

Tabla 21
Cantidad de Estudiantes por Clúster de la Sección B

Clúster	Cantidad de Estudiantes
CLUSTER 1	32
CLUSTER 2	12

Tabla 22
Estilos de aprendizaje del Clúster 1

CLUSTER 1		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Sensorial	Se orienta hacia hechos Percepción y datos concretos.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

Tabla 23
Estilos de aprendizaje del Clúster 2

CLUSTER 2		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Intuitiva	Prefiere lo conceptual y teórico a lo concreto.
Entrada	Visual	Elige ver lo que está aprendiendo por medio de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

SECCIÓN C

Tabla 24
Cantidad de Estudiantes por clúster de la Sección C

Clúster	Cantidad de Estudiantes
CLUSTER 1	17
CLUSTER 2	22

Tabla 25
Estilos de aprendizaje del Clúster 1

CLUSTER 1		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Sensorial	Se orienta hacia hechos Percepción y datos concretos.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

Tabla 26
Estilos de aprendizaje del Clúster 2

CLUSTER 2		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Intuitiva	Prefiere lo conceptual y teórico a lo concreto.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

SECCIÓN D

Tabla 27
Cantidad de estudiantes por clúster de la Sección D

clúster	Cantidad de Estudiantes
CLUSTER 1	17
CLUSTER 2	22

Tabla 28
Estilos de aprendizaje del Clúster 1

CLUSTER 1		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Sensorial	Se orienta hacia hechos Percepción y datos concretos.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

Tabla 29
Estilos de aprendizaje del Clúster 2

CLUSTER 2		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Intuitiva	Prefiere lo conceptual y teórico a lo concreto.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

SECCIÓN F

Tabla 30
Cantidad de estudiantes por clúster de la Sección F

Cluster	Cantidad de Estudiantes
CLUSTER 1	24
CLUSTER 2	17

Tabla 31
Estilos de aprendizaje del Clúster 1

CLUSTER 1		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Sensorial	Se orienta hacia hechos Percepción y datos concretos.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

Tabla 32
Estilos de aprendizaje del Clúster 2

CLUSTER 2		
DIMENSIÓN	**ESTILO**	**Características**
Procesamiento	Activo	Tiende a hacerlo mejor cuando puede trabajar en ello y en realidad experimenta o manipula cosas manualmente.
Percepción	Intuitiva	Prefiere lo conceptual y teórico a lo concreto.
Entrada	Visual	Prefiere ver lo que está aprendiendo a través de gráficos, diagramas e imágenes.
Comprensión	Secuencial	Prefiere disponer la información de forma lineal y ordenada

Resultados de Evaluaciones parciales en las secciones x asignatura

Tabla 33
Resultados de evaluación parciales

Ciclo	Asignatura	Sección	Promedio EP
I	DESARROLLO PERSONAL Y SOCIAL	A	13
I	FUNDAMENTOS DE PROGRAMACIÓN	A	10
I	INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	A	13
I	MATEMÁTICA BÁSICA	A	14
I	PROCESOS DE LA COMUNICACIÓN	A	13
I	DESARROLLO PERSONAL Y SOCIAL	B	15
I	FUNDAMENTOS DE PROGRAMACIÓN	B	8
I	INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	B	12
I	MATEMÁTICA BÁSICA	B	15
I	PROCESOS DE LA COMUNICACIÓN	B	14
I	DESARROLLO PERSONAL Y SOCIAL	C	15
I	FUNDAMENTOS DE PROGRAMACIÓN	C	12
I	INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	C	13
I	MATEMÁTICA BÁSICA	C	12
I	PROCESOS DE LA COMUNICACIÓN	C	14
I	DESARROLLO PERSONAL Y SOCIAL	D	15
I	FUNDAMENTOS DE PROGRAMACIÓN	D	12

I	INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	D	11
I	MATEMÁTICA BÁSICA	D	10
I	PROCESOS DE LA COMUNICACIÓN	D	11
I	DESARROLLO PERSONAL Y SOCIAL	F	14
I	FUNDAMENTOS DE PROGRAMACIÓN	F	14
I	INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	F	12
I	MATEMÁTICA BÁSICA	F	12
I	PROCESOS DE LA COMUNICACIÓN	F	14

Resultados de porcentaje por estado de aprobación en Examen Parcial

Tabla 34

Porcentaje de aprobados y desaprobados por sección

SECCIÓN	APROBADO	DESAPROBADO
A	64.06%	35.94%
B	60.64%	39.36%
C	60.33%	39.67%
D	51.77%	48.23%
F	64.60%	35.40%
Total general	**60.38%**	**39.62%**

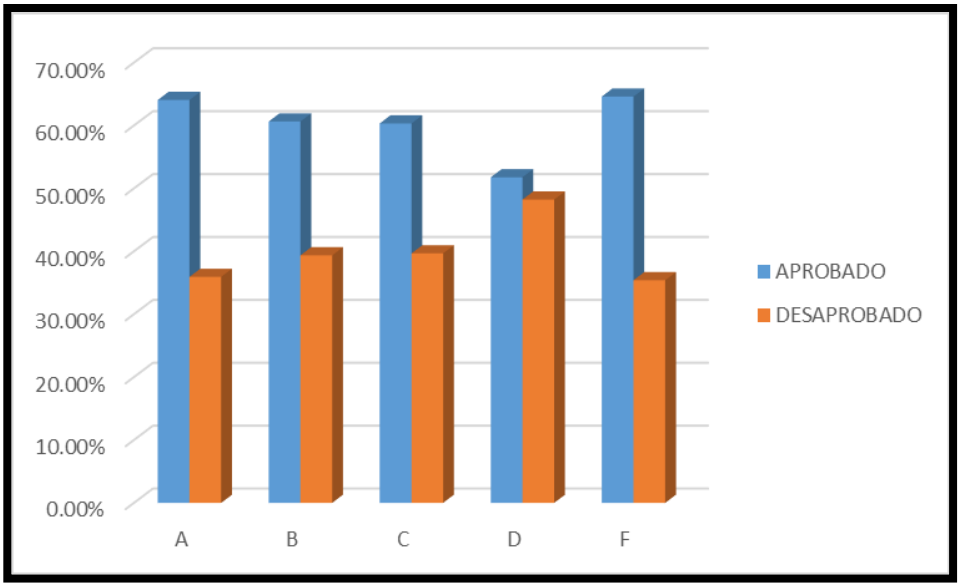


Figura 18. Diagrama de alumnos aprobados y desaprobados por sección

Tabla 35

Porcentaje de alumnos aprobados y desaprobados por curso

SECCIÓN	APROBADO	DESAPROBADO
A	64.06%	35.94%
DESARROLLO PERSONAL Y SOCIAL	81.13%	18.87%
FUNDAMENTOS DE PROGRAMACIÓN	29.41%	70.59%
INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	61.22%	38.78%
MATEMÁTICA BÁSICA	73.58%	26.42%
PROCESOS DE LA COMUNICACIÓN	74.00%	26.00%
B	**60.64%**	**39.36%**
DESARROLLO PERSONAL Y SOCIAL	84.00%	16.00%
FUNDAMENTOS DE PROGRAMACIÓN	14.00%	86.00%
INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	48.00%	52.00%
MATEMÁTICA BÁSICA	80.00%	20.00%
PROCESOS DE LA COMUNICACIÓN	77.55%	22.45%
C	**60.33%**	**39.67%**
DESARROLLO PERSONAL Y SOCIAL	82.98%	17.02%
FUNDAMENTOS DE PROGRAMACIÓN	56.86%	43.14%
INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	65.96%	34.04%
MATEMÁTICA BÁSICA	23.53%	76.47%
PROCESOS DE LA COMUNICACIÓN	76.09%	23.91%
D	**51.77%**	**48.23%**
DESARROLLO PERSONAL Y SOCIAL	84.09%	15.91%
FUNDAMENTOS DE PROGRAMACIÓN	57.69%	42.31%
INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	53.33%	46.67%
MATEMÁTICA BÁSICA	25.58%	74.42%
PROCESOS DE LA COMUNICACIÓN	35.71%	64.29%
F	**64.60%**	**35.40%**
DESARROLLO PERSONAL Y SOCIAL	76.09%	23.91%
FUNDAMENTOS DE PROGRAMACIÓN	71.74%	28.26%
INTRODUCCIÓN A LA INGENIERÍA DE SISTEMAS	47.73%	52.27%
MATEMÁTICA BÁSICA	48.89%	51.11%
PROCESOS DE LA COMUNICACIÓN	77.78%	22.22%
Total general	**60.38%**	**39.62%**

Resultados de la preprueba y la posprueba

Tabla 36
Datos de la Preprueba

KPI_01_PRE	KPI_02_PRE
10	7200
12	7800
8	6900
10	6600
10	7440
12	6720
15	8400
10	7560
12	7260
10	7980
20	7680
15	7740
14	6840
19	7560
14	7740
16	7080
15	8760
17	9120
20	6960
15	7080
10	7680
13	6840
17	8160
16	7620
10	9180
15	10560
14	8280
15	7500
20	11100
14	7800

Tabla 37
Datos de la postprueba

KPI_01_POST	KPI_02_POST
0.3	0.20
0.5	0.30
0.4	0.15
0.5	0.23
0.3	0.25
0.2	0.30
0.3	0.25
0.4	0.28
0.5	0.27
0.3	0.25
0.3	0.21
0.4	0.26
0.5	0.22
0.5	0.24
0.3	0.21
0.2	0.23
0.3	0.26
0.5	0.21
0.3	0.27
0.5	0.25
0.4	0.21
0.2	0.24
0.3	0.26
0.4	0.24
0.5	0.22
0.2	0.21
0.3	0.26
0.5	0.28
0.3	0.21
0.2	0.22

Tabla 38
Promedio de los indicadores de la preprueba y Post Prueba

Indicador	preprueba (media: x1)	postprueba (media: x2)	Comentarios
KPI_1: Tiempo empleado para determinar los clústeres de estilos de aprendizaje	13.9 horas	0.36 segundos	
KPI2: Tiempo empleado para determinar los estilos de aprendizaje por cluster	7200 minutos	0.20 minutos	

4.4. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

A. Indicador Tiempo empleado para determinar los clústeres de estilos de aprendizaje por sección**: KPI_1**

Tabla 39
Resultados de Pre-Prueba y Post Prueba para el KPI1

	Pre-prueba	Post-Prueba		
	10	0.3	0.3	0.3
	12	0.5	0.5	0.5
	8	0.4	0.4	0.4
	10	0.5	0.5	0.5
	10	0.3	0.3	0.3
	12	0.2	0.2	0.2
	15	0.3	0.3	0.3
	10	0.4	0.4	0.4
	12	0.5	0.5	0.5
	10	0.3	0.3	0.3
	20	0.3	0.3	0.3
	15	0.4	0.4	0.4
	14	0.5	0.5	0.5
	19	0.5	0.5	0.5
	14	0.3	0.3	0.3
	16	0.2	0.2	0.2
	15	0.3	0.3	0.3
	17	0.5	0.5	0.5
	20	0.3	0.3	0.3

	15	0.5	0.5	0.5
	10	0.4	0.4	0.4
	13	0.2	0.2	0.2
	17	0.3	0.3	0.3
	16	0.4	0.4	0.4
	10	0.5	0.5	0.5
	15	0.2	0.2	0.2
	14	0.3	0.3	0.3
	15	0.5	0.5	0.5
	20	0.3	0.3	0.3
	14	0.2	0.2	0.2
Promedio	13.9333333		0.36	
Meta Planteada		0.5		
Nº menor al Promedio		16	21	30
% menor al Promedio		53%	70%	

El 53 % del t**iempo empleado para determinar los clústeres de estilos de aprendizaje** en la Post-Prueba fueron menores que su tiempo promedio.

El 70 % del tiempo empleado **para determinar los clústeres de estilos de aprendizaje** en la Post-Prueba fueron menores que la meta planteada.

B. Indicador Tiempo empleado para determinar los estilos de aprendizaje por cluster**: KPI_2**

Tabla 40
Resultados de Pre-Prueba y Post-Prueba para el KPI2

	Pre-prueba	Post-Prueba		
	7200	0.20	0.20	0.20
	7800	0.30	0.30	0.30
	6900	0.15	0.15	0.15
	6600	0.23	0.23	0.23
	7440	0.25	0.25	0.25
	6720	0.30	0.30	0.30
	8400	0.25	0.25	0.25
	7560	0.28	0.28	0.28

	7260	0.27	0.27	0.27
	7980	0.25	0.25	0.25
	7680	0.21	0.21	0.21
	7740	0.26	0.26	0.26
	6840	0.22	0.22	0.22
	7560	0.24	0.24	0.24
	7740	0.21	0.21	0.21
	7080	0.23	0.23	0.23
	8760	0.26	0.26	0.26
	9120	0.21	0.21	0.21
	6960	0.27	0.27	0.27
	7080	0.25	0.25	0.25
	7680	0.21	0.21	0.21
	6840	0.24	0.24	0.24
	8160	0.26	0.26	0.26
	7620	0.24	0.24	0.24
	9180	0.22	0.22	0.22
	10560	0.21	0.21	0.21
	8280	0.26	0.26	0.26
	7500	0.28	0.28	0.28
	11100	0.21	0.21	0.21
	7800	0.22	0.22	0.22
Promedio	7838.00		0.24	
Meta Planteada		0.3		
Nº menor al Promedio		30	30	30
% menor al Promedio		100%	100%	100%

- El 100 % del Tiempo empleado para determinar los estilos de aprendizaje por clúster en la Post-Prueba fueron menores que su tiempo promedio.

- El 100 % Tiempo empleado para determinar los estilos de aprendizaje por clúster en la Post-Prueba fueron menores que la meta planteada.

4.5. CONTRASTACIÓN DE LA HIPÓTESIS

A continuación, se presentan las medias de los KPIs para la Pre-Prueba y Post-Prueba: Resultados numéricos.

Tabla 41
Contratación de la Hipótesis

Indicador	Pre-Prueba (Media: $\overline{X}_1$)	Post-Prueba (Media: $\overline{X}_2$)	Comentario
KPI$_1$: Tiempo empleado para determinar los clústeres de estilos de aprendizaje	13.9	0.36	
KPI2: Tiempo empleado para determinar los estilos de aprendizaje por cluster	7838.00	0.24	

Contrastación para el indicador: Tiempo empleado para determinar los clústeres de estilos de aprendizaje

Se debe validar el impacto que tiene el uso de un Sistema en el Tiempo empleado para determinar los clústeres de estilos de aprendizaje, llevado a cabo en la muestra. Se realiza una medición antes de utilizar el Sistema (Pre-Prueba) y otra después de utilizar el Sistema (Post-Prueba).

Hi: El uso de minería de datos disminuye el Tiempo empleado para determinar los clústeres de estilos de aprendizaje (Post-Prueba) con respecto a la muestra a la que no se aplicó (Pre-Prueba).

Solución:

a) Planteamiento de la Hipótesis

μ_1= Media del Tiempo empleado para determinar los clústeres de estilos de aprendizaje

Pre-Prueba.

μ_2= Media del Tiempo empleado para determinar los clústeres de estilos de aprendizaje

Post-Prueba.

$$H_0: \mu_1 \leq \mu_2$$

$$H_a: \mu_1 > \mu_2$$

b) Criterios de Decisión

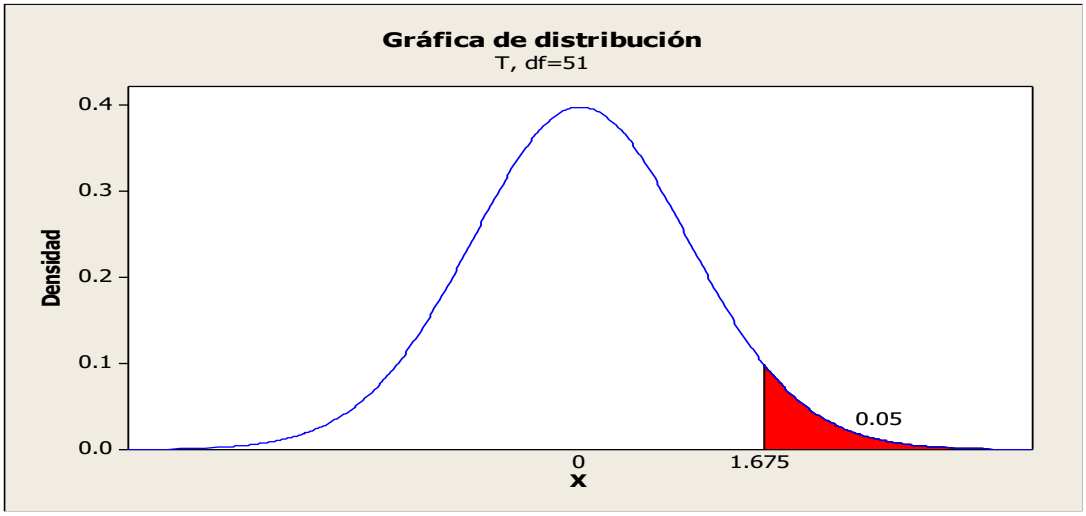


Figura 19. Distribución de probabilidades

c) Cálculo: Prueba t para prueba de medias de las dos muestras:

Tabla 42
Resumen de prueba t de student del KPI1

	Pre-Prueba	**Post-Prueba**
Media(x)	13.9	0.36
Desviación Estándar(S)	1.14	0.77
Observaciones	30	30
Diferencia hipotética de las medias	0	
t_calculado: t_c	**26.27**	
p-valor (una cola)	**0.000**	
Valor crítico de $t_{\alpha/2}$ (una cola)	**1.675**	

c) Decisión Estadística

Puesto que el valor-p=0.000 < α=0.05, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H_o), y la hipótesis alterna (H_a) es cierta. La prueba resultó ser significativa.

Contrastación para el indicador: Tiempo empleado para determinar los estilos de aprendizaje por cluster

Se debe validar el impacto que tiene el uso de un Sistema en el Tiempo empleado para determinar los estilos de aprendizaje por cluster, llevado a cabo en la muestra. Se realiza una medición antes de utilizar el Sistema (Pre-Prueba) y otra después de utilizar el Sistema (Post-Prueba).

Hi: El uso de un Sistema KDD disminuye el Tiempo empleado para determinar los estilos de aprendizaje por cluster (Post-Prueba) con respecto a la muestra a la que no se aplicó (Pre-Prueba).

Solución:

a) Planteamiento de la Hipótesis

μ_1= Media del Tiempo empleado para determinar los estilos de aprendizaje por cluster Pre-Prueba.

μ_2= Media del Tiempo empleado para determinar los estilos de aprendizaje por cluster Post-Prueba.

$$H_0: \mu_1 \leq \mu_2$$

$$H_a: \mu_1 > \mu_2$$

b) Criterios de Decisión

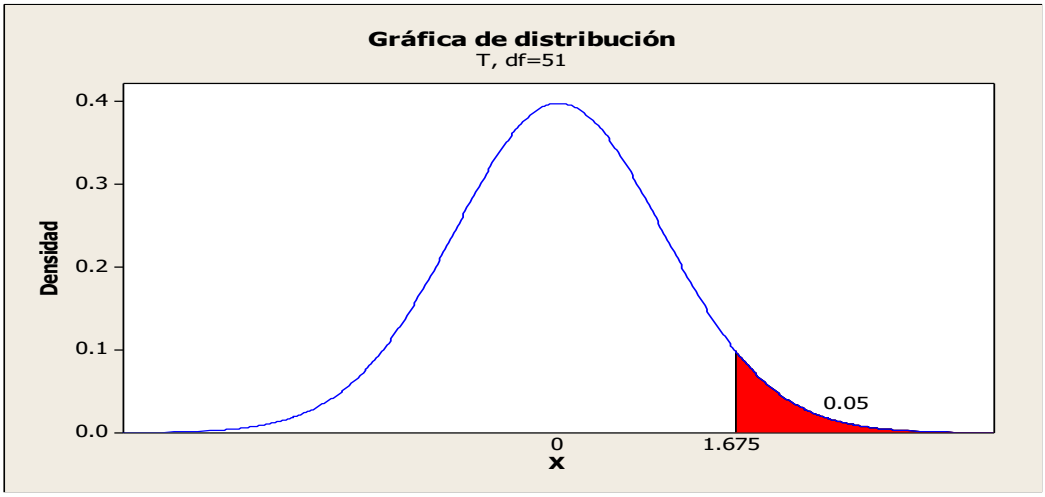


Figura 20. Distribución de Probabilidad del KPI1

c) Cálculo: Prueba t para prueba de medias de las dos muestras:

Tabla 43
Resumen de prueba t de student del KPI1

	Pre-Prueba	**Post-Prueba**
Media(x)	7838	0.2397
Desviación Estándar(S)	1048	0.0325
Observaciones	30	30
Diferencia hipotética de las medias	0	
t_calculado: t_c	40.97	
p-valor (una cola)	**0.000**	
Valor crítico de $t_{\alpha/2}$ (una cola)	**1.675**	

d) Decisión Estadística

Puesto que el valor-p=0.000 < α=0.05, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H_o), y la hipótesis alterna (H_a) es cierta. La prueba resultó ser significativa.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

a. Se logró mejorar el Tiempo empleado para determinar los clústeres de estilos de aprendizaje de 13.9 horas (836 minutos) paso a 0.36 minutos; esto supone una mejora de más del 99.6%.

b. Se logró mejorar Tiempo empleado para determinar los estilos de aprendizaje por clúster, de 7200 minutos a 0.20 minutos lo que da una mejora de más del 99.9%

c. La técnica de clúster permitió identificar los 2 clúster predominantes en cada sección para luego ser identificado el estilo de aprendizaje predominante.

5.2. RECOMENDACIONES

a) Para adaptar o ajustar la didáctica, se hace primordial modelar al estudiante según las cambiantes a las cuales se quiere adaptar la enseñanza: estilos de estudio, opciones, interacción, etc. La personalización de la didáctica en sí misma constituye la formación centrada en el alumno, por consiguiente, consigue una particular consideración el hecho de poder modelar a cada estudiante, comprender sus aspectos cognitivos, y en funcionalidad de esto, planear el acto formativo.

b) Para adaptar o ajustar la didáctica, se hace primordial modelar al estudiante según las cambiantes a las cuales se quiere adaptar la enseñanza: estilos de estudio, opciones, interacción, etc. La personalización de la didáctica en sí misma constituye la formación centrada en el alumno, por consiguiente, consigue una particular consideración el hecho de poder modelar a cada estudiante, comprender sus aspectos cognitivos, y en funcionalidad de esto, planear el acto formativo

REFERENCIAS BIBLIOGRÁFICAS

TESIS

Andrés, J. (2009). *Minería de Datos Aplicada al Análisis de la Deserción en la carrera de Analista en Sistemas de Computación* (Tesis de pre grado). Recuperado de http://88.25.6.252/files/1496984174_TFAGermanPAUTSCHFinal.pdf

Conde, O. (2015). Relación entre Estilos de Aprendizaje Autoestima y Rendimiento Escolar en Estudiantes del Primer Año de Bachillerato (Tesis de pre grado). Recuperado de https://es.calameo.com/read/000501353ca90c71fe1fc

Fischer, E. (2012). *Modelo para la Automatización del Proceso de Determinación de Riesgo de Deserción en Estudiantes Universitarios* (Tesis de maestría). Recuperado de http://repositorio.uchile.cl/bitstream/handle/2250/111188/cf-fischer_ea.pdf?sequence=1&isAllowed=y

Gonzáles, J. (2013). *Sistema de Apoyo para la Acreditación de la Calidad de Programas Académicos de la Universidad de Caldas, Aplicando Técnicas en Minería de Datos* (Tesis de pre grado). Recuperado de http://repositorio.autonoma.edu.co/jspui/bitstream/11182/350/1/Msc.GyDlloSoftInformeFinal_JuanCarlosGonzalez.pdf

LIBROS

Gil, G. (2009). *Data Mininig - Minería De Datos y SQL*. Recuperado de http://sbiblio.uandina.edu.pe/cgi-bin/koha/opac-detail.pl?biblionumber=1359

Vieria, L., Ortiz, L., y Ramírez, S. (2009). *Introducción a la Minería De Datos*. Recuperado de https://books.google.com.pe/books/about/Introducci%C3%B3n_a_la_Miner%C3%ADa_de_Datos.html?id=jIJEhHyESFsC

Artículos

Bonilla, A., Ojeda., M y Echeverria, F. (2006). Implementación de Minería de Datos Basado en Redes Bayesianas para la Toma De Decisiones en los Registros Académicos. *Iberoamericana de Educación.* 42(1), 180-|190. Recuperado de https://www.researchgate.net/publication/28791884_Implementacion_De_Mineria_De_Datos_Basada_En_Redes_Bayesianas_Para_La_Toma_De_Decisiones_En_Los_Registros_Academicos

Durán, E y Costaguata, R. (2007). Aplicación de Técnicas de Minería de Datos para Descubrir Estilos de Aprendizaje. *Iberoamericana de Educación*, 42(2), 5-9. Recuperado de https://www.researchgate.net/publication/28150021_Mineria_de_datos_para_descubrir_estilos_de_aprendizaje

Martí, J. (2009). Estilos de Aprendizaje y Modelo de Usuario en los Sistemas Hipermedia Adaptativos. *Universidad de la Habana*, 12(1), 1-10, Recuperado de https://www.monografias.com/trabajos-pdf4/estilos-aprendizaje-y-modelo-usuario/estilos-aprendizaje-y-modelo-usuario.pdf

Ramirez, Y. y Orteaga, J. (2012). Diagnóstico del Estilo de Aprendizaje Predominante Basado en Minería de Datos el Modelo de Felder: Aplicaciones Al Elearnig 3.0. *Dialnet*, 12(14), 14-20, Recuperado de https://dialnet.unirioja.es/servlet/articulo?codigo=4651452

Riquelme, J., Ruiz, R. y Gilbert, K. (2006). Minería de Datos: Conceptos y Tendencias, *Redalyc*. 10(29), 11-18. Recuperado de http://www.redalyc.org/articulo.oa?id=92502902

Sitios web

Alonso, C, Gallego, D y HOney, P. (1995). Los estilos de aprendizaje. Procedimientos de diagnóstico y mejora. Recuperado de https://periplosenred.blogspot.com/2011/08/4-estilos-de-aprendizaje-segun-alonso.html

Felder, R y Silverman, L. (1988). Estilos de aprendizaje y de enseñanza en la educación de ingeniería. Recuperado de http://piagetanos.blogspot.com/p/inicio.html

Garcia, J. (2009). Instrumento de medición de aprendizaje. Recuperado de https://www.researchgate.net/publication/38290638_Instrumentos_de_medicion_de_Estilos_de_Aprendizaje

IBM (2012). Manual CRISP-DM de IBM SPSS Modeler. Recuperado de ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/15.0/es/CRISP-DM.pdf

Universidad Autónoma del Perú (2018). Universidad Autónoma del Perú. Recuperado de https://www.google.com/maps/place/Universidad+Aut%C3%B3noma+del+Per%C3%BA/@-12.1955204,-76.9739164,17z/data=!4m5!3m4!1s0x9105b9989f7875ef:0xac24a8fceeb849a!8m2!3d-12.1955257!4d-76.9717277

Webmining Consultores (2011). KDD:Proceso de Extradicción de Conocimiento. Recuperado de http://www.webmining.cl/2011/01/proceso-de-extraccion-de-conocimiento/

ANEXOS

TEST DE FELDER

NOMBRE: _____

CARRERA: _____

CICLO: _____

SECCIÓN: _____

DISTRITO DE RESIDENCIA: _____

TIPO DE COLEGIO: PARTICULAR ☐ ESTATAL ☐

EDAD: _____

SEXO: _____

INSTRUCCIONES

Encierre en un círculo la opción "a" o "b" para indicar su respuesta a cada pregunta. Por favor seleccione solamente una respuesta para cada pregunta.

Si tanto "a" y "b" parecen aplicarse a usted, seleccione aquella que se aplique más frecuentemente.

Entiendo mejor algo

a) si lo practico

b) si pienso en ello

Me considero

a) realista

b) innovador

3. Cuando pienso acerca de lo que hice ayer, es más probable que lo haga en base

a) una imagen

b) palabras

4. Tengo tendencia a entender

a) los detalles de un tema pero no ver claramente su estructura completa

b) la estructura completa pero no ver claramente los detalles

5. Cuando estoy aprendiendo algo nuevo, me ayuda

a) hablar de ello

b) pensar en ello

6. Si yo fuera profesor, yo preferiría dar un curso

a) que trate sobre hechos y situaciones reales de la vida

b) que trate con ideas y teorías

7. Prefiero obtener información nueva de

a) imágenes, diagramas, gráficos o mapas

b) instrucciones escritas o información verbal

8. Una vez que entiendo

a) todas las partes, entiendo el total

b) el total de algo, entiendo como encajan sus partes

9. En un grupo de estudio que trabaja con un material difícil, es más probable que

a) participe y contribuya con ideas

b) no participe y solo escuche

10. Es más fácil para mí

a) aprender hechos

b) aprender conceptos

11 En un libro con muchas imágenes y gráficas es más probable que

a) revise cuidadosamente las imágenes y las gráficas

b) me concentre en el texto escrito

12. Cuando resuelvo problemas de matemáticas

a) generalmente trabajo sobre las soluciones con un paso a la vez

b) frecuentemente sé cuáles son las soluciones, pero luego tengo dificultad para imaginarme los pasos para llegar a ellas

13. En las clases a las que he asistido

a) he llegado a saber cómo son muchos de los estudiantes

b) raramente he llegado a saber cómo son muchos estudiantes

14. Cuando leo temas que no son de ficción, prefiero

a) algo que me enseñe nuevos hechos o me diga cómo hacer algo

b) algo que me dé nuevas ideas en que pensar

15. Me gustan los profesores

a) que utilizan muchos esquemas en el pizarrón

b) que toman mucho tiempo para explicar

16. Cuando estoy analizando un cuento o una novela

a) pienso en los incidentes y trato de acomodarlos para configurar los temas

b) me doy cuenta de cuáles son los temas cuando termino de leer y luego tengo que regresar y encontrar los incidentes que los demuestran

17. Cuando comienzo a resolver un problema de tarea, es más probable que

a) comience a trabajar en su solución inmediatamente

b) primero trate de entender completamente el problema

18. Prefiero la idea de

a) certeza

b) teoría

19. Recuerdo mejor

a) lo que veo

b) lo que oigo

20. Es más importante para mí que un profesor

a) exponga el material en pasos secuenciales claros

b) me dé un panorama general y relacione el material con otros temas

21. Prefiero estudiar

a) en un grupo de estudio

b) solo

22. Me considero

a) cuidadoso en los detalles de mi trabajo

b) creativo en la forma en la que hago mi trabajo

23. Cuando alguien me da direcciones de nuevos lugares, prefiero

a) un mapa

b) instrucciones escritas

24. Aprendo

a) a un paso constante. Si estudio con ahínco consigo lo que deseo

b) en inicios y pausas. Me llego a confundir y súbitamente lo entiendo

25. Prefiero primero

a) hacer algo y ver que sucede

b) pensar como voy a hacer algo

26. Cuando leo por diversión, me gustan los escritores que

a) dicen claramente lo que desean dar a entender

b) dicen las cosas en forma creativa e interesante

27. Cuando veo un esquema o bosquejo en clase, es más probable que recuerde

a) la imagen

b) lo que el profesor dijo acerca de ella

28. Cuando me enfrento a un cuerpo de información

a) me concentro en los detalles y pierdo de vista el total de la misma

b) trato de entender el todo antes de ir a los detalles

29. Recuerdo más fácilmente

a) algo que he hecho

b) algo en lo que he pensado mucho

30. Cuando tengo que hacer un trabajo, prefiero

a) dominar una forma de hacerlo

b) intentar nuevas formas de hacerlo

31. Cuando alguien me enseña datos, prefiero

a) gráficos

b) resúmenes con texto

32. Cuando escribo un trabajo, es más probable que lo haga (piense o escriba)

a) desde el principio y avance

b) en diferentes partes y luego las ordene

33. Cuando tengo que trabajar en un proyecto de grupo, primero quiero realizar

a) una " tormenta de ideas " donde cada uno contribuye con ideas

b) la "tormenta de ideas" en forma personal y luego juntarme con el grupo para comparar las ideas

34. Considero que es mejor elogio llamar a alguien
a) sensible
b) imaginativo

35. Cuando conozco gente en una fiesta, es más probable que recuerde
a) cómo es su apariencia
b) lo que dicen de sí mismos

36. Cuando estoy aprendiendo un tema, prefiero
a) mantenerme concentrado en ese tema, aprendiendo lo más que pueda de él
b) hacer conexiones entre ese tema y temas relacionados

37. Me considero
a) abierto
b) reservado

38. Prefiero cursos que dan más importancia a
a) material concreto (hechos, datos)
b) material abstracto (conceptos, teorías)

39. Para divertirme, prefiero
a) ver televisión
b) leer un libro

40. Algunos profesores inician sus clases haciendo un bosquejo de lo que enseñarán. Esos bosquejos son
a) algo útiles para mí
b) muy útiles para mí

41. La idea de hacer una tarea en grupo con una sola calificación para todos
a) me parece bien
b) no me parece bien

42. Cuando hago grandes cálculos
a) tiendo a repetir todos mis pasos y revisar cuidadosamente mi trabajo

b) me cansa hacer su revisión y tengo que esforzarme para hacerlo

43. Tiendo a recordar lugares en los que he estado
a) fácilmente y con bastante exactitud
b) con dificultad y sin mucho detalle

44. Cuando resuelvo problemas en grupo, es más probable que yo
a) piense en los pasos para la solución de los problemas
b) piense en las posibles consecuencias o aplicaciones de la solución en un amplio rango de campos

GLOSARIO DE TÉRMINOS

ALGORITMO: El término algoritmo es un procedimiento o fórmula para resolver un problema

BASES DE DATOS: (DataBase). Conjunto de datos relacionados que se almacenan de forma que se pueda acceder a ellos de manera sencilla, con la posibilidad de relacionarlos, ordenarlos en base a diferentes criterios, etc. Las bases de datos son uno de los grupos de aplicaciones de productividad personal más extendidos. Entre las más conocidas pueden citarse dBase, Paradox, Access y Aproach, para entornos PC, y Oracle, ADABAS, DB/2, Informix o Ingres, para sistemas medios y grandes.

CLUSTERING (**AGRUPAMIENTO**). Proceso de dividir un conjunto de datos en grupos mutuamente excluyentes de tal manera que cada miembro de un grupo esté lo "más cercano" posible a otro, y grupos diferentes estén lo "más lejos" posible uno del otro, donde la distancia está medida con respecto a todas las variables disponibles.

DATA MINING: La extracción de información predecible escondida en grandes bases de datos.

HTML: es el acrónimo inglés de HyperText Markup Language, que se traduce al español como Lenguaje de Marcas Hipertextuales. Es un lenguaje de marcación diseñado para estructurar textos y presentarlos en forma de hipertexto, que es el formato estándar de las páginas web. Gracias a Internet y a los navegadores como Internet Explorer, Opera, Firefox, Netscape o Safari, el HTML se ha convertido en uno de los formatos más populares y fáciles de aprender que existen para la elaboración de documentos para web.

MINERIA DE DATOS: Es una disciplina de la informática, que surge de la necesidad de desarrollar herramientas y modelos de análisis para grandes y complejos volúmenes de datos con el objetivo de extraer conocimiento, y forma parte de un proceso llamado Descubrimiento de Conocimiento en Bases de Datos Knowledge Discovery in Databases (KDD) definido como: todo el proceso no trivial de identificar patrones en datos que sean válidos, novedosos, potencialmente útiles y

humanamente comprensibles”, con el objeto de predecir de forma automatizada tendencias y comportamientos y descubrir modelos previamente desconocidos.

URL: significa Uniform Resource Locator, es decir, localizador uniforme de recurso. Es una secuencia de caracteres, de acuerdo a un formato estándar, que se usa para nombrar recursos, como documentos e imágenes en Internet, por su localización.