

UNIVERSIDAD DE LOS ANDES FACULTAD DE INGENIERIA
DIVISION DE ESTUDIOS DE POSTGRADO
PROGRAMA DE ESTUDIOS DE POSTGRADO EN COMPUTACION
MERIDA – VENEZUELA



**La analítica social de aprendizaje en los ciclos autonómicos de tareas de
análisis de datos para aulas inteligentes**

Autor: Ing. Omar Buendía

Tutores: Dr. José L. Aguilar C.

Dra. Tania Rodríguez.

Trabajo de Grado presentado ante la ilustre Universidad de Los Andes como requisito para optar al grado de *Magister Scientiae en Computación*.

Mérida Marzo de 2018

La analítica social de aprendizaje en los ciclos autonómicos de tareas de análisis de datos para aulas inteligentes.

Ing. Omar R. Buendía.

Proyecto de Grado -- Postgrado en Computación, 88 Páginas.

Escuela de Ingeniería de Sistemas, Universidad de Los Andes, 2018.

Resumen: la Analítica Social de Aprendizaje (SLA, por sus siglas en inglés) se deriva de la analítica de aprendizaje (Learning Analytics, LA), y en ambos casos se busca obtener información oculta en grandes cantidades de datos, normalmente de carácter educativo. Ahora bien, la SLA se enfoca principalmente en el análisis de las redes sociales (Social Network Analysis, SNA) y de la web, para tratar de descubrir patrones de interacción y de comportamiento de los actores sociales educativos. Esta investigación, busca incorporar la SLA en los ambientes inteligentes educativos, en especial, en las aulas inteligentes (SaCI). Para llevar a cabo este trabajo, se hará un análisis de los datos externos procedentes de la web y de las redes sociales más utilizadas por los estudiantes, para especificar tareas de analítica de datos sobre ellos, tales como clasificación, agrupación y búsqueda de patrones, con el fin de mejorar el proceso de enseñanza-aprendizaje que se da en SaCI. En general, dichas tareas de análisis de datos se organizarán en ciclos autonómicos, de manera de integrarlas entre sí. Además, el ciclo autonómico permitirá automatizar la ejecución de esas tareas y la generación de conocimiento, de tal manera de supervisar permanentemente el proceso de enseñanza-aprendizaje en SaCI, observándolo, analizándolo y planificando las acciones que permitan optimizarlo. Para el desarrollo de las tareas de SLA, se analizará el

uso de la Web Semántica, Minería Semántica, Minería de Grafos, Big Data y Linked Data, entre otras técnicas. Finalmente, se especificará la arquitectura computacional de los ciclos autonómicos de tareas de SLA y LA para SaCI.

Palabras claves: análisis de datos, aulas inteligentes, analítica social de aprendizaje, computación autonómica.

INDICE

INDICE	4
Capítulo 1	1
Introducción	1
1.1 Generalidades	1
1.2 Planteamiento del Problema	2
1.3 Alcance	3
1.4 Objetivos	4
1.4.1 Objetivo General	4
1.4.2 Objetivo Específicos	4
1.5 Antecedentes	4
1.6 Estructura del documento	7
Capítulo 2	8
Marco Teórico	8
2.1 Analítica de Datos (AdD)	8
2.1.1 Objetivos Fundamentales de AdD.	9
2.1.2 Proceso de AdD	9
2.2 Analítica de Aprendizaje	9
2.2.1 Características de LA	10
2.2.2 Métodos de análisis de LA	10
2.2.3 Analítica de Aprendizaje Social	11
2.3 Minería Semántica	12
2.3.1 Minería de datos semántica (MDS)	13
2.3.2 Minería Web Semántica (MWS)	13
2.3.3 Minería Ontológica	13
2.3.4. Minería de texto	14
2.4 MIDANO versión extendida	16
2.4.1 Fase 1: Identificación de fuentes para la extracción de conocimiento en una organización	18
2.4.2 Fase 2: Preparación y tratamiento de los Datos	20
2.4.3 Fase 3 Desarrollo de las tareas de AdD	23
2.5 Computación Autónoma	24
2.5.1 Características básicas de un sistema de computación autónoma	24
2.5.2 Computación Autónoma basada en la arquitectura MAPE-K	26
2.5.3 Ciclos Autónomos de Tareas de Análisis de Datos	27
2.6 Salón de Clase Inteligente (SaCI)	28
Capítulo 3	31
Especificación de los Ciclos Autónomos	31
3.1 Ciclo 1: Determinar el paradigma de aprendizaje adecuado para un curso.	31
3.2 Ciclo 2: Determinar el recurso educativo ideal para un estudiante	36
3.3 Ciclo 3: Identificar a los estudiantes con necesidades específicas	39

3.4 Ciclo 4: Evitar la deserción estudiantil	42
Capítulo 4	46
Desarrollo del Ciclo Autonomico “Definir el paradigma de aprendizaje adecuado para un Curso”	46
4.1 Caracterización General del Ciclo Autonomico	46
4.2 Especificación del Ciclo Autonomico	47
4.2.1 Representación Gráfica del Ciclo Autonomico.	47
4.2.2 Tareas que componen el Ciclo Autonomico.	47
4.2.3 Descripción de las tareas del ciclo autonomico actual	48
4.2.4 Macro-Algoritmo del Ciclo Autonomico de Tareas	50
4.3 Incorporación del CA en el PTL de SaCI.	52
Capítulo 5	53
Escenarios Experimentales y Análisis de Resultados para el Ciclo Autonomico “Definir el paradigma de aprendizaje adecuado para un Curso”	53
5.1 Escenarios Experimentales para el Ciclo Autonomico	53
5.1.1 Escenario de prueba Nro. 1.	53
5.1.2 Escenario de prueba Nro. 2.	64
5.2 Análisis de Resultados	68
5.2.1 Resultados Escenario 1	69
5.2.2 Resultados Escenario 2	70
5.2.2 Comparación con trabajos similares	71
Capítulo 6	73
Conclusiones y Recomendaciones	73
6.1 Conclusiones	73
6.2 Recomendaciones	75
Referencias Bibliográficas	75
Apéndices	78

INDICE DE TABLAS

	Páginas
Tabla 2.1. Preguntas y ejemplos para determinar los elementos MIDANO.	19
Tabla 3.1. Desarrollo de tareas de analítica de datos en el ciclo 1.	33
Tabla 3.2. Descripción de tareas en el Ciclo 2.	37
Tabla 3.3. Descripción de las tareas del Ciclo 3.	41
Tabla 3.4. Descripción de las tareas del Ciclo 4.	44
Tabla 4.1. Descripción de la tarea 1.	48
Tabla 4.2. Descripción de la tarea 2.	49
Tabla 4.3. Descripción de la tarea 3.	50
Tabla 4.4. Descripción de la tarea 4.	50
Tabla 4.5. Macro – Algoritmo del Ciclo Autónomo.	51
Tabla 5.1. Macro-Algoritmo del Ciclo categoría 5284.	54
Tabla 5.2. Tablas y campos utilizados en las consultas SQL.	56
Tabla 5.3. Tablas y campos utilizados en las consultas SQL.	61
Tabla 5.4. Conteo de Frecuencias de las actividades general.	62
Tabla 5.5. Macro-Algoritmo del Ciclo categoría 5223.	64
Tabla 5.6. Tablas y campos utilizados en las consultas SQL.	67
Tabla 5.7. Tablas y campos utilizados en las consultas SQL.	71

INDICE DE FIGURAS

	Páginas
Figura 2.1. Modelo general de minería de texto.	15
Figura 2.2. Fases de MIDANO Extendido.	17
Figura 2.3. Aspectos que conforman cada etapa de las fases de la metodología.	17
Figura 2.4. Etapas que conforman la fase 1.	18
Figura 2.5: Etapas que conforman la fase 2.	20
Figura 2.6: Etapas que conforman la fase 3.	23
Figura 2.7: Arquitectura de la Computación Autonomica.	25
Figura 2.8: Ciclo MAPE_K.	26
Figura 2.9: Ciclo Autonomico de Tareas.	27
Figura 2.10: Middleware AmICL.	29
Figura 3.1: Esquema conceptual del ciclo 1.	33
Figura 3.2: Modelo de datos del primer ciclo autonomico.	35
Figura 3.3: Esquema conceptual de Ciclo 2.	38
Figure 3.4: Esquema conceptual del Ciclo 3.	41
Figura 3.5: Esquema conceptual del Ciclo 4.	44
Figura 4.1: Representación gráfica del ciclo autonomico 1.	47
Figura 4.2: Conversación del PTL SaCI usando el servicio provisto por el CA.	52
Figura 5.1: Consulta a la base de datos categoría 5284.	55
Figura 5.2: Agrupamiento K-means del curso de computación .	57
Figura 5.3: Data frame del timeLine de los estudiantes.	58
Figura 5.4: Extracción de tweeds del archivo estudiantesCorpus.csv.	59
Figura 5.5: Creación del vector y la matriz de frecuencias de actividades.	59

Figura 5.6: Filtro de las actividades de aprendizaje.	60
Figura 5.7: Frecuencia de actividades de un estudiante.	60
Figura 5.8: Consulta de las actividades Moodle.	62
Figura 5.9: Estilo de aprendizaje de la categoría de Computación UTPL.	65
Figura 5.10: Consulta a la base de datos categoría de Derecho.	67

Capítulo 1

Introducción

1.1 Generalidades

La inteligencia Ambiental (AmI, por sus siglas en inglés) es un área de la computación que se enfoca en tener espacios tecnológicamente enriquecidos, que eficientemente apoyen a las personas en su vida diaria [1]. Entre las principales características que encontramos en la AmI se pueden mencionar, la de detectar información del entorno y razonar con los datos acumulados, de tal manera de seleccionar las acciones adecuadas, con el fin de beneficiar a los usuarios [1]. Algunos campos de aplicación de la AmI son: las ciudades inteligentes, las casas inteligentes, los hospitales inteligentes, los espacios de trabajo inteligentes, y los ambientes inteligentes educativos. En este trabajo nos concentraremos en estos últimos, específicamente, en las aulas inteligentes (también conocidos como salones de clase inteligentes, SaCI) [2].

En general, en un AmI se generan y acumulan muchos datos, los cuales deben ser analizados con el fin de descubrir el conocimiento implícito en ellos. La analítica de datos es la ciencia que examina los datos en bruto, con el propósito de descubrir conocimiento para sacar conclusiones [16]. La Analítica de aprendizaje (LA, por sus siglas en inglés) es un subcampo de la analítica de datos, que busca extraer conocimiento desde grandes volúmenes de datos educativos, con el propósito de mejorar el proceso de enseñanza–aprendizaje. Por otra parte, la SLA es un subcampo de la LA, que busca estudiar los datos exógenos al proceso educativo, apoyándose en la Web Semántica, la Linked Data, entre otros conceptos, con el fin de descubrir conocimiento oculto para mejorar los procesos de

enseñanza–aprendizaje, como por ejemplo, los patrones comportamentales de los estudiantes en las redes sociales [9].

Finalmente, el concepto de ciclo autónomo de tareas de análisis de datos ha sido introducido, con el fin de integrar y automatizar las actividades de analítica de datos, de tal manera de permanente descubrir y utilizar el conocimiento en los procesos organizacionales [8].

En este trabajo se analizará el uso de las tareas de SLA en un SaCI, y se especificará la arquitectura computacional que soporte las tareas de SLA en el mismo. La arquitectura definirá las tareas genéricas fundamentales de SLA en un SaCI, y, la forma de integración de las diferentes técnicas de SLA, basada en el concepto de ciclo autónomo.

1.2 Planteamiento del Problema

En la actualidad se están generando grandes cantidades de datos los cuales no son explotados para extraer conocimiento útil para mejorar los procesos de enseñanza-aprendizaje. Para la gestión y uso de esos datos masivos, se requieren de nuevos paradigmas desde las ciencias computacionales, tales como Big Data, Minería Semántica, Ciencia de los Datos, entre otros [16]. En particular, para analizar los datos provenientes de fuera de los entornos organizacionales, que implica manipular grandes volúmenes de datos, ha aparecido una nueva área en las ciencias computacionales, denominada análisis social de datos [9]. La analítica social de datos se ha usado de manera pionera en los procesos de marketing, explotando como fuentes de datos masivos a las redes sociales. Un tipo particular de analítica social de datos es la analítica de aprendizaje social (SLA), la cual estudia los datos fuera de los entornos educativos, con el objetivo de mejorar los procesos de enseñanza-aprendizaje.

Este trabajo se propone analizar la incorporación de la SLA en SaCI. En los SaCIs se generan muchos datos e información desde sus dinámicas internas (ambientes virtuales de aprendizaje, sistemas académicos, sistemas de recomendación, entre otros). Por lo tanto, las

tareas asociadas al análisis de los mismos son requeridas, de tal manera de crear mecanismos para explotarlos de formas automática y con la mínima intervención humana posible, con el fin de generar conocimiento.

El problema radica en que si no se tiene cuidado con esta labor, la generación del conocimiento puede ser estéril y no útil para el proceso de enseñanza-aprendizaje que se da en SaCI. En ese sentido, en [8] se ha desarrollado el concepto de ciclo autónomo de tareas de análisis de datos, con el fin de integrar y automatizar las tareas de análisis de datos, de tal manera de supervisar, analizar y optimizar permanentemente los AmIs.

Así, este trabajo estudia el problema de la definición e integración de las tareas de SLA en un SaCI, para analizar los datos exógenos del AmI, y generar conocimiento útil para el proceso de enseñanza-aprendizaje que se da en él. En particular, para resolver el problema de integración y automatización de las tareas de SLA en un SaCI, se usará el concepto de ciclo autónomo [8]. En ese sentido, este trabajo especificará la arquitectura computacional que permita el despliegue de ciclos autónomos de tareas de SLA en un SaCI, integrando para ello, las técnicas provenientes de las áreas de Big Data, Minería Semántica, Ciencia de los Datos, entre otras. Además, definirá las tareas genéricas fundamentales de SLA, para soportar los procesos de enseñanza-aprendizaje en SaCI, las cuales se deberán integrar de manera armónica con las tareas de LA presentes actualmente en los ciclos autónomos de SaCI [8].

1.3 Alcance

En este trabajo, se presenta el diseño de los iniciales Ciclos Autónomos para SaCI, haciendo hincapié en sus tareas de SLA. Además, se presenta un ejemplo de implementación de uno de ellos, y sus posibles usos en el contexto de SaCI. También, se explica cómo sería integrado dicho Ciclo Autónomo de Tareas de SLA en la arquitectura actual de SaCI, en particular a través de las conversaciones de sus agentes en el Proceso de Tutoría en Línea (PTL), pero en específico, invocadas por ellos como servicios que requieren para sus procesos de toma de decisión. Por ejemplo, esto le va a permitir a

agentes de SaCI, como el de Entorno Virtual de Aprendizaje (VLE) y el Sistema Recomendador (SR), gestionar las actividades académicas adecuadas para cada perfil de estudiantes de un curso dado, en el contexto del Salón de Clase Inteligente (SaCI).

1.4 Objetivos

1.4.1 Objetivo General

Incorporar la Analítica Social de Aprendizaje (SLA) en SaCI para mejorar sus procesos de enseñanza-aprendizaje, usando el concepto de ciclos autonómicos.

1.4.2 Objetivo Específicos

- Caracterizar las fuentes de datos sociales exógenas a un SaCI, de utilidad para mejorar sus procesos de enseñanza-aprendizaje.
- Analizar los conceptos y herramientas de las áreas de Big Data, Minería Semántica, Minería de Grafos, entre otras, para la extracción de conocimiento útil desde datos sociales.
- Especificar tareas de SLA desde esos datos exógenos, que generen conocimiento útil para mejorar los procesos de enseñanza-aprendizaje en SaCI.
- Diseñar ciclos autonómicos que integren y automaticen el uso de las tareas de SLA en SaCI.
- Especificar la plataforma computacional, que permita el despliegue de esas tareas de SLA en SaCI.
- Implementar un prototipo de la plataforma computacional, y de algunas de las tareas genéricas de SLA requeridas por SaCI.
- Desarrollar experimentos sobre SaCI, para evaluar el impacto de la introducción de tareas de SLA en los procesos de enseñanza-aprendizaje de SaCI.

1.5 Antecedentes

En el mundo de hoy, las iniciativas de automatización en la educación han demostrado ser

una herramienta y un recurso valioso, por lo que las investigaciones en el tema han tenido un auge importante [2]. Actualmente, dichas iniciativas se han venido desplegando en áreas tales como de aprendizaje electrónico (e-learning), Computación en la nube (cloud computing), AmI, entre otras [2].

En específico, para el ámbito educativo se han desarrollado una gran cantidad de productos de software, entre los que tenemos [2]: Sistemas de Tutoría Inteligente (Intelligent Tutorial Systems (ITS)), Sistemas de Aprendizaje Colaborativo Soportado por Computadora (Computer-Supported Collaborative Learning Systems (CSCL)), Sistemas de Gestión de Contenido (Content Management Systems, (CMS)), Sistemas de Gestión de Aprendizaje (Learning Management Systems, (LMS)), Entornos Virtuales de Aprendizaje (Virtual Learning Environments (VLE)), entre otros.

Este trabajo estará enfocado específicamente a los Salones de Clase Inteligentes (SaCIs), por lo que a continuación se presentan algunos antecedentes de relevancia en ese contexto [2].

Uno de los trabajos a mencionar es el proyecto ISABEL [3], el cual es un sistema de e-learning basado en sistemas multiagentes. En ese proyecto, los estudiantes se dividen en grupos con perfiles similares, y cada grupo es gestionado por un agente tutor. Adicionalmente, hay un agente profesor asociado a cada grupo de e-learning, el cual provee recomendaciones a los estudiantes y está siempre activo en ese sitio.

Otro proyecto interesante es el proyecto de salones de clases inteligentes [4], que propone un salón de clases interactivo en tiempo real, que incorpora tecnologías de computación de gran impacto dentro del aprendizaje a distancia. El objetivo de este proyecto es integrar las experiencias de los profesores con las tecnologías, en la tele-educación y en los salones de clase tradicionales. La iniciativa propone tener estudiantes remotos en un salón de clases aumentado, de manera que el profesor pueda interactuar con un estudiante remoto, tal como lo hace con un estudiante presencial.

Por otro lado, el proyecto AARTIC [5] propone un ambiente inteligente que da soporte a estudiantes de ingeniería del software en sus asignaciones. El sistema ayuda a los

estudiantes a entender los conceptos usados. El proyecto propone dos agentes, el primero de ellos supervisa las actividades de los estudiantes en el sistema, y el segundo permite al profesor evaluar las clases como un todo. En [6] presentan un modelo de ambientes educativos inteligentes, así como una arquitectura conceptual para sistemas que soporten dicho modelo.

En [2] explica que un Ambiente Inteligente (AmI) en la educación debe integrar diferentes aspectos relacionados con el proceso de enseñanza, como el perfil de los estudiantes, el paradigma pedagógico utilizado, entre otras cosas, en tiempo real. En ese trabajo, ellos caracterizan un SaCI teniendo en cuenta esos aspectos, mediante el paradigma de sistemas multiagentes. En particular, ellos definen los diferentes componentes de SaCI (hardware y software), sus propiedades, y dan ejemplos de las conversaciones presentes en él. Además, especifican un medio de gestión de servicios (middleware) para la comunidad de agentes de SaCI, que se denomina AmICL [7], el cual explota la computación en la nube, para brindar servicios educativos al entorno de aprendizaje.

En cuanto a SLA, en [12] presenta algunas aproximaciones, todas enfocadas en la forma en la que los estudiantes construyen conocimiento juntos, en su contexto cultural y social. De acuerdo con [12], SLA es un proceso en el cual los estudiantes no trabajan solos, en vez de eso, participan en actividades sociales, integrándose unos con otros en el aprendizaje. El trabajo anterior da una idea de la variedad de investigaciones que se pueden realizar en SLA, muestra cómo el conocimiento generado a partir del comportamiento social, puede ser usado para resolver problemas educativos, apoyar la toma de decisiones en este contexto, y al mismo tiempo, genera nuevos retos de investigación. Uno de ellos es un objetivo de este trabajo, ¿Cómo pueden automatizarse mediante ciclos autónomos las tareas de Social Learning Analytics en los Salones de Clase Inteligentes, para mejorar el proceso de enseñanza y aprendizaje?

En [13] hacen una presentación de la idea de automatización de los ciclos autónomos basados en diferentes contextos, tales como, ciclos autónomos para el modelo educativo basados en el paradigma de la nube, también ciclos autónomos basados en el paradigma eco-conectivista, y ciclos autónomos en el contexto de los salones de clase inteligentes

SaCI. Más adelante, en [8] son especificados con más detalle los ciclos autonómicos, además, varios objetivos vitales de SaCI son automatizados mediante el uso de ciclos autonómicos.

Finalmente, [14] propone un ciclo autonómico más elaborado, en el que incorpora el concepto de competencias como servicio en el ciclo autonómico. En ese trabajo se sugiere incluir el manejo de las competencias de los estudiantes en un proceso autonómico, mediante la incorporación de una tarea que da este servicio informativo a las demás. Esta tarea especial crea el perfil de competencias de cada estudiante, y determina la trayectoria de aprendizaje adecuada a él.

1.6 Estructura del documento

Este documento está organizado de la siguiente manera: el Capítulo 2 formula las bases teóricas necesarias para la correcta comprensión de los conceptos utilizados durante el trabajo. El Capítulo 3 describe la especificación del ciclo autonómico, el diseño de las tareas de analítica de datos y analítica de datos social que lo conforman, y la definición del modelo de datos que serán utilizados en la investigación. El Capítulo 4 muestra la implementación del ciclo autonómico, presenta la descripción de cada una de las tareas que lo conforman, y el macro algoritmo de cada una de ellas. El capítulo 5 presenta la experimentación y el análisis de los resultados obtenidos. Finalmente, el capítulo 6, presenta las conclusiones de la investigación, así como las recomendaciones para los trabajos futuros.

Capítulo 2

Marco Teórico

En este capítulo se presentan las bases teóricas para el entendimiento de este trabajo, comenzaremos con la definición de Analítica de datos, a continuación se presentan los conceptos relacionados a la minería semántica, los Salones de Clase Inteligentes (SaCI) y los Ciclos Autónomos de Tareas de análisis de datos (CA).

2.1 Analítica de Datos (AdD)

La Analítica de Datos AdD es la ciencia que examina los datos en bruto con el propósito de buscar conocimiento implícito en ellos, sacar conclusiones, y generar información que a simple vista no es evidente. Actualmente se usa en muchos ámbitos, como por ejemplo:

- En la agricultura, para obtener los patrones de cultivos de plantas, patrones de comportamientos de fertilizantes y plaguicidas etc.
- La economía, para establecer patrones del comportamiento en las bolsas de valor y mercados económicos actuales.

En general, la AdD es usada en la actualidad en las áreas de las ciencias que involucren el manejo de volúmenes de datos de gran tamaño. En este sentido, las organizaciones han tenido que centrarse en la explotación de los datos para obtener una ventaja competitiva, todo esto ha sido posible debido a que: las computadoras de hoy en día son más poderosas, el trabajo en red y la interconexión está presente en muchos aspectos de la vida diaria, se han desarrollado algoritmos que puedan conectar conjuntos de datos, la web ha evolucionado de tal manera que ha contribuido en el aumento del volumen de datos, entre otras cosas.

2.1.1 Objetivos Fundamentales de AdD.

- Ayudar a solucionar los problemas de las organizaciones desde la perspectiva de los datos.
- Comprender los principios de extracción y conocimiento útil de datos.

2.1.2 Proceso de AdD

Debido a que el éxito de AdD sólo puede medirse en términos de lo bien que ayudan a solucionarse los problemas de las organizaciones desde las perspectivas de los datos, el proceso debe:

- Identificar los objetivos de la organización.
- Recoger los datos necesarios para medir sus objetivos.
- Analizar los datos.
- Sacar conclusiones basadas en los datos.

2.2 Analítica de Aprendizaje

La Analítica de aprendizaje (LA, por sus siglas en inglés *Learning Analytics*) es una extensión de la AdD, la cual consiste en el uso de datos provenientes de contextos educativos tales como, universidades, los estudiantes, sistemas educativos, para construir modelos o patrones para mejorar el entorno en el cual se produce el proceso de aprendizaje. Generalmente, LA se ocupa del desarrollo de métodos que aprovechan conjuntos de datos educativos para apoyar el proceso de aprendizaje. Este uso de los datos puede extraer conocimiento de base de datos educativa, o en nuestro caso, de salones de clases inteligentes, para comprender mejor el proceso de aprendizaje que se da en él. Por ejemplo, un Entorno Virtual de Aprendizaje (VLE, por sus siglas en inglés), permite la generación de grandes cantidades de datos relacionados con los procesos de enseñanza-aprendizaje, lo que ofrece la posibilidad de extraer información valiosa que puede emplearse para mejorar el rendimiento de los estudiantes.

2.2.1 Características de LA

- Se emplea en organizaciones educativas, para mejorar los cursos actuales o desarrollar nuevas ofertas de currículo.
- Es usada por administradores de instituciones educativas, para tomar decisiones sobre asuntos tales como marketing, reclutamiento, problemas de desempeño, entre otros.
- Es usada por los estudiantes, para mejorar sus procesos de aprendizaje, patrones de aprendizaje, logros, con materiales personalizados, recursos educativos adecuados a sus perfiles, entre otros.
- Es usada por los profesores, para identificar estudiantes de "riesgo" de abandono escolar o fracaso del curso, predecir los estudiantes que requieren apoyo y atención extra, etc.

2.2.2 Métodos de análisis de LA

Para [15], los métodos de análisis de LA incluyen:

- **El Análisis de Contenido:** usado para analizar los recursos que los estudiantes crean, por ejemplo sus ensayos.
- **Análisis del discurso:** tiene como objetivo capturar datos significativos sobre las interacciones de los estudiantes. Se diferencia de Social Learning Analytics en que se enfoca en el estudio del lenguaje utilizado, y el comportamiento de los estudiantes, en el aula de clase, y no analiza sus iteraciones en una red social.
- **Análisis de Redes Sociales (SNA):** cuyo objetivo es explorar el papel de la interacción social en el aprendizaje, la importancia de las redes de aprendizaje.
- **Analítica de disposición:** que trata de capturar los datos relativos a las disposiciones de los estudiantes sobre su propio aprendizaje, la relación de ellos con su aprendizaje. Por ejemplo, los alumnos "curiosos" pueden estar más inclinados a hacer preguntas, y estos datos pueden ser capturados y analizados para el análisis de aprendizaje.

2.2.3 *Analítica de Aprendizaje Social*

Es una sub área de LA (SLA por sus siglas en inglés *Social Learning Analytics*), la cual se enfoca en cómo los estudiantes construyen conocimiento en un contexto social, a menudo sin ser supervisados, en colaboración con otros. La SLA centra la atención en los elementos de aprendizaje que son relevantes cuando se está aprendiendo en una cultura participativa. A continuación se presentan algunas características relevantes de SLA:

- Los usuarios almacenan conjuntos de datos y crean representaciones visuales.
- Los conjuntos de datos y representaciones son accesibles para otros usuarios de la red o sitio web.
- Los usuarios pueden crear nuevas e interesantes representaciones, así como comentarios asociados a las intervenciones en el grupo.
- Se pueden usar un blogs y wikis para conducir esta inteligencia social.

En general, es posible recuperar los datos sociales de una variedad de fuentes sociales, tales como Twitter, Facebook, Instagram, Wikipedia, YouTube, entre otros. Este tipo de información se puede obtener a través de una interfaz de programación de aplicaciones (API, por sus siglas en inglés), que es un conjunto de subrutinas métodos y procedimientos usados en la actualidad por las aplicaciones en general, para dar acceso a los datos en sus plataformas.

2.2.3.1 Aspectos a considerar en la SLA

Análisis de datos sofisticados: El análisis de datos sociales debe tomar en consideración una serie de factores (contexto, el contenido, el sentimiento) para proporcionar información adicional.

La consideración del tiempo: Lo más relevante de un día (o incluso una hora) puede no ser en el siguiente. Ser capaz de ejecutar con rapidez el análisis es imperativo.

Análisis de la influencia: la comprensión del impacto potencial de individuos específicos puede ser clave en la comprensión de cómo los mensajes podrían estar resonando. No se trata sólo de la cantidad, también tiene mucho que ver con la calidad.

Análisis de las Redes: los datos sociales migran, crecen (o mueren) en base a cómo se propagan a través de la red. Es como una actividad viral, que se inicia y se propaga.

2.2.3.2 Elementos relevantes del aprendizaje social

- SLA busca hacer visible, comportamientos y patrones en el ambiente de aprendizaje.
- Los alumnos no son solitarios y hacen actividades sociales, ya sea interactuando con los demás, o dejando trazas en las plataformas de sus actividades.
- Los logros individuales se transmiten a través de la interacción y la colaboración.
- El éxito es una combinación de los conocimientos, habilidades individuales, el medio ambiente, el uso de herramientas, y la capacidad de trabajar juntos.
- La comprensión del proceso de aprendizaje obliga a prestar atención a los procesos del grupo en la construcción del conocimiento.
- La atención debe centrarse no sólo en los alumnos, sino también en las herramientas y contextos.

2.3 Minería Semántica

Uno de los problemas más importantes y difíciles en la minería de datos, es la incorporación del conocimiento del dominio. Cuando el conocimiento del dominio está disponible, vale la pena explotarlo para generar nuevo conocimiento. Ese proceso es conocido como Minería Semántica.

La minería semántica se encarga de extraer conocimiento desde diferentes fuentes semánticas, tales como: páginas web, contenido sin estructura en la web, contenido estructurado en la web, grafos anotados, ontologías, tabla de datos, entre otros.

La Minería Semántica se divide en tres grandes grupos, la Minería de Datos Semántica (MDS), la Minería de la Web Semántica (MWS) y la Minería Ontológica (MO). Una descripción breve de cada una de ellas se hace a continuación:

2.3.1 Minería de datos semántica (MDS)

Su objetivo es incorporar contenido semántico a los datos. Para ello, realiza un proceso de enriquecimiento semántico usando ontologías o fuentes semánticas, que son alineadas o mapeadas a los datos para enriquecerlos. También, otro proceso que se puede hacer con las técnicas de Minería de Datos es la identificación de patrones en los datos, que permiten aportar un nuevo conocimiento al contexto.

2.3.2 Minería Web Semántica (MWS)

Integra la Web Semántica y la Minería Web, para explotar el contenido de la Web Semántica, expresado en lenguajes como OWL, RDF, entre otros. La Web es minada, para extraer el conocimiento escondido en ella. Para ello, la minería Web aplica técnicas de la minería de datos (como por ejemplo, las técnicas de Minería de Texto) para extraer, por ejemplo, patrones en la Web. Algunas de las tareas de extracción de conocimiento de la MWS es sobre el uso de la Web, sobre su contenido, o sobre su estructura. El minado sobre el contenido básicamente es una forma de Minería de Texto, que se aplica al contenido desplegado en la Web. El minado de la estructura estudia el patrón de enlace de las páginas web, usando técnicas de análisis y Minería de Grafos. Finalmente, el minado del uso de la Web se enfoca en minar el historial de uso de los usuarios, para descubrir los patrones de comportamiento en las consultas que hacen a una página, los movimientos que se hacen entre páginas, entre otras cosas.

2.3.3 Minería Ontológica

De acuerdo con [16], la minería ontológica hace referencia a la extracción de patrones de comportamiento y de conocimiento, además de otras características, desde ontologías. Las ontologías son estructuras que expresan un esquema formal dentro de un dominio dado. [16] propone que una ontología “es una especificación explícita de una conceptualización”, y además, señala que el conocimiento en las ontologías se formaliza a través de seis componentes: clases, atributos, relaciones, funciones, axiomas e instancias. Algunos autores agrupan esas técnicas en el área de ingeniería ontológica, que se refiere al conjunto de actividades que tienen que ver con el proceso de desarrollo de ontologías, los métodos y metodologías para construir ontologías, y las herramientas y lenguajes que las soportan.

2.3.3.1 Aprendizaje de Ontologías (AO)

Construir y ensamblar ontologías ha sido un cuello de botella en el diseño para el ingeniero de conocimiento, debido a los conceptos asociados a la semántica del dominio. El proceso de AO tiene que ver con la extracción de elementos ontológicos a partir de diferentes fuentes, que dan como resultado la construcción o actualización de una ontología existente [16]. Los elementos ontológicos que normalmente se extraen son conceptos, palabras, relaciones, funciones, axiomas, e instancias. Como el objetivo es extraer conceptos y relaciones relevantes de una fuente de datos, que puede ser estructurada, semi-estructurada, o no estructurada, es necesario utilizar diferentes técnicas de recuperación de información, tales como:

- **Aprendizaje de ontologías a partir de textos:** En este enfoque se aprenden conceptos y relaciones utilizando textos. El objetivo es extraer conocimiento desde textos
- **Aprendizaje de ontología desde instancias. Instancias particulares,** por ejemplo, aquellas que son tomadas de una página web, de un texto libre.
- **Aprendizaje de Ontologías desde esquemas:** esquemas de bases de datos, modelos entidad-relación, esquemas XML, etc., son usados para generar ontologías por procesos de re-ingeniería. Por ejemplo, la generación de ontologías a través de mapeos, es decir, usando XML u otro lenguaje declarativo, el cual describe el emparejamiento entre un esquema de base de datos y una ontología.

2.3.4. Minería de texto

La minería de texto es una área de la ciencia de los datos, que intenta resolver problemas de procesamiento textual, con la ayuda de técnicas de la minería de datos, como por ejemplo: aprendizaje automático, procesamiento del lenguaje natural (PNL), Recuperación de Información (IR) y Gestión del conocimiento. La minería de Texto también es definida en [16], como un proceso en que los usuarios interactúan con una colección de documentos a lo largo del tiempo, utilizando una serie de herramientas de análisis para obtener información. La minería de textos incluye la recopilación y pre-procesamiento de documentos, categorización de texto, extracción de información, extracción de términos, la

generación de representaciones intermedias como: distribución de términos, agrupación de términos, análisis de tendencias, reglas de asociación, etc., y visualizaciones de los resultados, que dan origen a conocimiento que no existía previamente en ningún texto de la colección, pero que surge al relacionar el contenido de algunos de ellos (Ver figura 2.1).

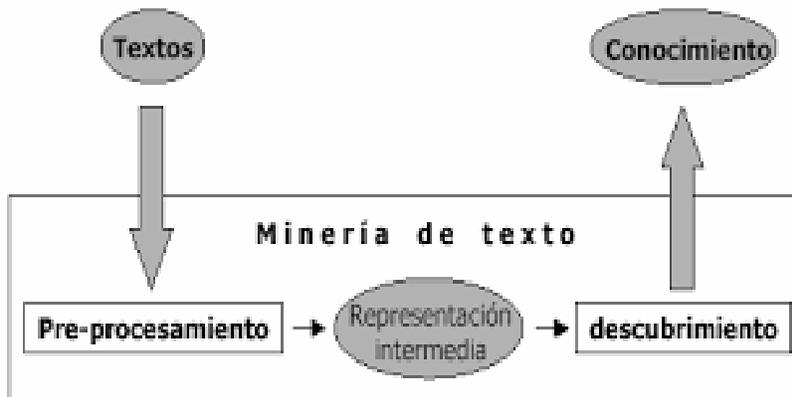


Figura 2.1: Modelo general de minería de texto.

2.3.4.1 Características de los modelos en Minería de Texto

En general, los modelos de representación en minería de texto tratan de cumplir con 2 objetivos fundamentales [16]:

- **Establecer el nivel semántico:** que busca las características para poder representar con un alto grado de precisión el nivel semántico de un documento, o un grupo de documentos. En este sentido, se realizan vocabularios controlados, fuentes de conocimiento como diccionarios ontológicos, o bases de conocimiento.
- **Identificar las características de forma computacional eficiente:** de forma que sean más ordenadas computacionalmente, para tratar de descubrir patrones. En este sentido se realizan validaciones, normalizaciones o referencias cruzadas de características.

Existen diferentes modelos para representar y organizar documentos, de manera de crear grupos más pequeños, con características semánticas más ricas. En general, de acuerdo con [19], los tres tipos de representación de documentos más utilizados son:

- **Caracteres:** las letras, los números, los caracteres especiales y los espacios, son el bloque de las características semánticas de alto nivel, tales como palabras, términos y conceptos. Una representación a nivel de caracteres puede incluir el conjunto completo de todos los caracteres de un documento, o algún subconjunto filtrado.
- **Palabras:** Las palabras específicas seleccionadas directamente de un documento "nativo", son lo que se puede describir como el nivel básico de riqueza semántica.
- **Términos:** Los términos son palabras simples y frases multipalabra seleccionadas directamente del corpus de un documento nativo, mediante metodologías de extracción de términos. Las características a nivel de término, en el sentido de esta definición, sólo pueden estar compuestas de palabras y expresiones específicas encontradas dentro del documento nativo que sean generalmente representativas.

2.4 MIDANO versión extendida

MIDANO es una metodología diseñada para el desarrollo de aplicaciones basadas en minería de datos, desarrollada en la Universidad de los Andes [20]. Está compuesta por tres grandes fases, como lo son: (1) Identificación de fuentes para la extracción de conocimiento en una organización, (2) Preparación y tratamiento de los Datos y (3) Desarrollo del ciclo o ciclos autonómicos. En la Figura 2.2 se muestra el flujo de desarrollo de dichas fases, en la imagen se observa que se puede retroceder a fases anteriores de ser necesario. Más detalles de esta metodología están en [20].



Figura 2.2: Fases de MIDANO Extendido fuente [20].

Cada fase de la metodología está concebida en etapas que se ejecutan secuencialmente mediante una serie de pasos. Los elementos que conforman cada etapa se muestran en la Figura 2.3

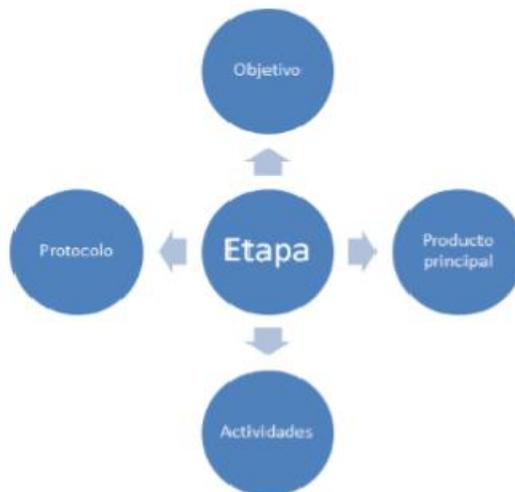


Figura 2.3: Aspectos que conforman cada etapa de las fases de la metodología.

- **Objetivo:** describe cual es la meta que se quiere cumplir en la etapa respectiva

- **Producto principal:** que es lo que se debe producir al final de la etapa.
- **Protocolo:** es el conjunto de procedimientos, preguntas o estudios que se deben realizar para desarrollar la etapa.
- **Actividades:** describe las tareas que se designan a los investigadores y a la organización/empresa para lograr el objetivo de la etapa.

2.4.1 Fase 1: Identificación de fuentes para la extracción de conocimiento en una organización

Esta fase tiene como finalidad realizar un proceso de ingeniería de conocimiento, orientado a organizaciones/empresas, de las cuales no se conoce o se tiene poca información de los problemas, o los procesos a estudiar (Ver Figura 2.4).



Figura 2.4: Etapas que conforman la fase 1.

Etapa 1: Conocimiento de la organización

Objetivo: Conocer la organización/empresa, sus objetivos, procesos, objetos y actores

Producto de la Fase: Documento con la descripción de los elementos que contiene por lo menos los siguientes ítems:

- Descripción de los elementos de la institución/empresa y sus características
- Descripción de las relaciones entre estos elementos
- Organización de estos elementos

Protocolo de la Fase: Familiarización con los procesos sobre los cuales se puede realizar la ingeniería de conocimiento

- Identificación de la fuente de conocimiento
- Familiarización con los ambientes computacionales donde se encuentran los datos a ser utilizados en cada proceso.

Para la descripción de cada elemento, se pueden realizar las preguntas dadas en la Tabla 2.1.

Elemento	Preguntas	Ejemplos
Objetivo	¿Cuál es la razón de ser de la institución?	Conocer, determinar, establecer, la finalidad de la institución/empresa.
procesos	¿Cuáles son las actividades que permiten alcanzar los objetivos de la institución?	Procesos de producción o administrativos.
Objetos	¿Qué cosas o entidades se manipulan en los procesos de la institución?	Pueden ser físicos o abstractos, departamentos, documentos, herramientas, plantas.
Actores	¿Quiénes ejecutan los procesos?	Personas, sistemas, máquinas, etc.

Tabla 2.1: Preguntas y ejemplos para determinar los elementos MIDANO [20].

Actividades de la fase

Actividad 1: Generar un documento que permita conocer la institución/empresa, respondiendo las interrogantes de la tabla 1.1.

Actividad 2: Estudio de la institución/empresa con la información proporcionada por la misma. Generar dudas sobre el funcionamiento de la institución/empresa, de sus procesos y objetivos expuestos en el documento consignado al grupo de investigación.

2.4.2 Fase 2: Preparación y tratamiento de los Datos

En esta fase se plantea realizar la preparación de los datos desarrollando dos etapas. Los productos más resaltantes de esta fase son las vistas minables (conceptual y operativa), las variables objetivos, y el modelo de datos multidimensional (Ver figura 2.5).



Figura 2.5: Etapas que conforman la fase 2.

Para aplicar AdD sobre un problema en específico, es necesario contar con un historial de datos asociado al problema en estudio. Para realizar tareas de AdD es necesario tener los datos integrados en una sola vista, la cual comúnmente se conoce como Vista Minable. Existen dos tipos de vista minable

- **Vista Minable Conceptual (VMC):** describe en detalle cada una de las variables a tomar en cuenta para c/tarea de AdD.
- **Vista Minable Operativa (VMO):** Es el resultado de cargar los datos del historial y de realizar la etapa de tratamiento de datos, basado en la información de la VMC. La VMO se traduce a lo que se conoce como Vista Minable en la literatura, para realizar tareas de MD.

Con esas vistas se construye el modelo de datos multidimensional.

Etapa 1: Definición de Modelos de Datos

Objetivos: En esta etapa se deben producir dos aspectos concretos, la vista minable conceptual y la vista minable operativa.

Productos principales

- Características de los repositorios donde se encuentran los datos
- Vista minable conceptual
- Vista minable operativa
- Descripción de las variables objetivos
- Organización de estos elementos en un modelo de datos multidimensional

Protocolo de la Fase: Familiarización con los modelos y estructuras de datos de la empresa:

- Comprender la fuente de datos de entrada
- Generar la VMC y la VMO
- Generar las tablas del modelo de datos multidimensional de cada Ciclo Autónomo

Actividades

Por parte de la institución/empresa:

- Proporcionar a los investigadores la información de los datos asociado al escenario futuro seleccionado

- Proveer los datos asociados a la VMC provenientes de los servidores de la institución/ empresa
- Proporcionar a los investigadores información que les permita definir las variables objetivo

Por parte de los ingenieros de conocimiento:

- Generar una descripción de los datos y las relaciones que tienen con las variables.
- Conocer cómo están almacenados los datos
- Seleccionar y ubicar las variables objetivos en la VMC
- Diseñar el Modelo de datos Multidimensional

Etapa 2: Tratamiento de los datos

Objetivos:

- Construir la tabla con las operaciones de Extracción, Transformación y Carga, (modelo ETL) para las variables identificadas en la VMC
- Cargar los datos

Protocolo de la etapa

- Integración de los datos de entrada a través de un almacén de datos.

Productos principales

- Tabla ETL

Actividades

Por parte de los investigadores:

Actividad: Realizar el proceso de limpieza, transformación y reducción de los datos en la VMO.

2.4.3 Fase 3 Desarrollo de las tareas de AdD

En esta fase se plantea realizar la preparación y el desarrollo de los ciclos autónomos de tareas y de cada una de las tareas de AdD (Ver figura 2.6).



Figura 2.6: Etapas que conforman la fase 3.

En esta etapa se evalúa la herramienta computacional utilizada para el desarrollo de las tareas de AdD, y se implementan los ciclos autónomos de tareas de analítica de datos.

Etapa 1: Especificación detallada de los requerimientos de la herramienta computacional

Objetivos: captar los requerimientos no funcionales.

Protocolo de la etapa: Requisitos de interfaz de usuario, Interfaces de software, Requerimientos de desempeño, Adicionalmente se pueden mencionar: de portabilidad, costos, rendimiento, accesibilidad, entre otros.

Productos principales: Informe de requerimientos no funcionales

Etapa 2: Especificación tecnológica del ciclo autónomo de Tareas de AdD

Objetivos:

- Caracterizar la implementación tecnológica del ciclo autónomo de tareas de AdD.

Protocolo de la etapa

- Escoger las técnicas de AdD para las tareas en el Ciclo Autonomico.
- Seleccionar el Software para realizar cada tarea de AdD
- Definir cuáles son los datos de entrenamiento y de prueba a usar
- Definir las interfaces entre las tareas del Ciclo Autonomico
- Definir una estrategia para la validación de las técnicas seleccionada (cruzada, etc.).

Productos principales

- Documento con la especificación tecnológica del ciclo

2.5 Computación Autonomica

El término computación autónoma fue propuesto en el año 2001 por Paul Horn en la Academia Nacional de Ingeniería de la Universidad de Harvard, basado en una relación entre el sistema nervioso autónomo y las capacidades del cuerpo para autorregularse. Paul Horn, de IBM, sugirió automatizar operaciones de cómputo sin intervención humana, basada en esa idea, para que se gobernaran de manera autónoma como lo hace el sistema nervioso, que protege nuestro cuerpo regulando funciones vitales como la respiración, la temperatura corporal y el pulso cardiaco, liberando al cerebro de esas tareas.

2.5.1 Características básicas de un sistema de computación autónoma

- Debe ser capaz de hacer un análisis continuo de sí mismo, sus conexiones, dispositivos y recursos.
- Debe ser capaz de configurarse y reconfigurarse dinámicamente, según las necesidades.
- Debe buscar siempre optimizar su rendimiento.
- Debe realizar tareas de auto-sanación para evitar situaciones disfuncionales, redistribuyendo recursos y reconfigurándose.

- Debe ser capaz de controlar su seguridad y protegerse de los ataques.
- Debe ser capaz de reconocer y adaptarse a las necesidades de los sistemas que coexisten en su entorno.

De acuerdo con [18], la computación autónoma define una arquitectura compuesta de 6 niveles (ver Figura. 2.7).

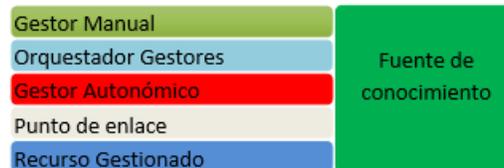


Figura 2.7: Arquitectura de la Computación Autónoma Fuente [18]

Recurso Gestionado: puede ser cualquier tipo de recurso (hardware o software) que puede ser gestionado. El recurso gestionado se controla a través de sus sensores y actuadores.

Punto de Enlace: Enlaza a los sensores y/o actuadores requeridos para la gestión de los recursos.

Gestor Autónomo: Implementa los lazos de control inteligentes que automatizan las tareas de autorregulación/autogestión de las aplicaciones.

Orquestador de Gestores Autónomos: Debido a que un sistema autónomo, puede contar con varios gestores autónomos que necesitan trabajar en conjunto, para garantizar el funcionamiento correcto de los recursos, este nivel proporciona el canal de comunicación para la coordinación entre ellos.

Manejador Manual: Permite a los humanos configurar los gestores autónomos para realizar su tarea de autogestión, proveyendo para esto una interfaz hombre-máquina que permite conectar al hombre con el gestor autónomo.

Fuentes de Conocimiento: Proporciona acceso a los conocimientos requeridos para la gestión autónoma del sistema.

2.5.2 Computación Autónoma basada en la arquitectura MAPE-K

La computación Autónoma basada en el modelo MAPE-K es un tipo de arquitectura que es capaz de almacenar el conocimiento para que sea usado en la toma de decisiones (ver Figura 2.8).



Figura 2.8: Ciclo MAPE_K.

Como se observa en la Fig. 2.2, el ciclo MAPE-K está formado por cuatro componentes:

Monitor: es el componente de supervisión del ciclo MAPE. Se encarga de la supervisión de las propiedades del entorno, de percibir las y organizarlas a través de sensores [17].

Analizador: se encarga de realizar una interpretación usando los datos que son enviados desde los sensores.

Planificador: toma en cuenta los eventos detectados por el analizador, para establecer una serie de pasos que serán enviados al ejecutor para ser procesados.

Ejecutor: es el componente encargado de ejecutar el plan. El ejecutor sólo recibe un mensaje del componente planificador, dicho mensaje contiene todas las acciones a realizar, en una secuencia específica.

2.5.3 Ciclos Autonomicos de Tareas de Análisis de Datos

En general, un “Ciclo Autonomico de Tareas de Analítica de Datos” puede definirse como un conjunto de tareas de AdD cuyo propósito es mejorar el proceso que estudia. Este conjunto de tareas, interactúa con otras, y tienen diferentes roles para llevar a cabo su proceso: **Observar** el proceso, **analizar e interpretar** lo que pasa en él, y **tomar decisiones** para mejorar el contexto donde es aplicado [8] (Ver figura 2.9).

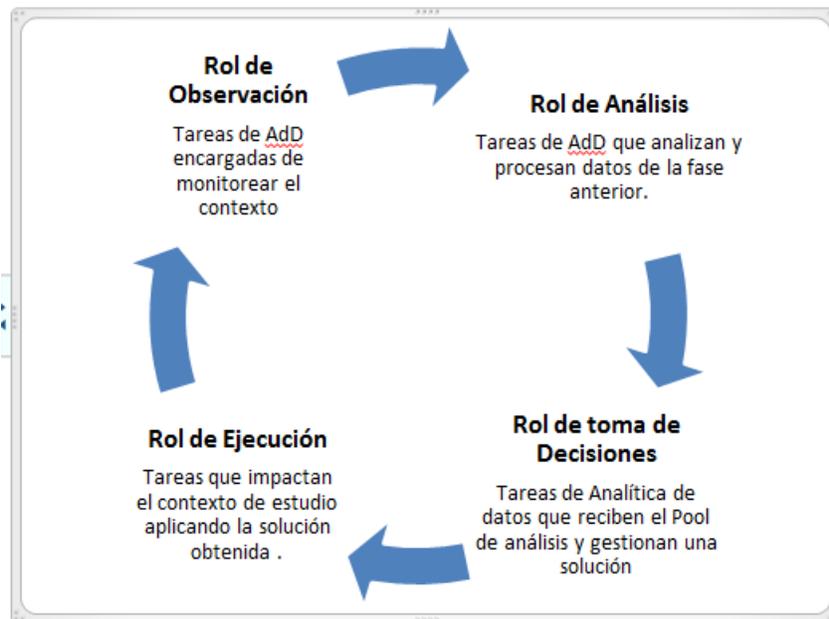


Figura 2.9: Ciclo Autonomico de Tareas.

Esta integración de tareas de analítica de datos en un ciclo cerrado, permite resolver problemas complejos, que hasta ahora no se habían podido estudiar por la cantidad de conocimiento requerido para la solución. En este sentido, es muy importante integrar las tareas de analítica de datos de forma coherente para generar conocimiento útil, estratégico para el logro de los objetivos.

El concepto del Ciclo, surge entonces, como una estructura circular, en las que son definidas las metodologías, herramientas y estrategias que permiten la integración de las tareas de AdD, en un complejo orden para la toma de decisiones. Los roles de cada tarea en el ciclo autónomo son:

Tareas de Observación: son un conjunto de tareas, encargadas de monitorear el proceso. Deben capturar datos e información acerca del comportamiento del entorno.

Tareas de Análisis del sistema: son un conjunto de tareas cuyo propósito es interpretar, entender y diagnosticar los datos dentro del contexto del ciclo. Esto significa que estas tareas construyen modelos de conocimiento acerca de las dinámicas en el ciclo, usando los datos provenientes de la etapa de observación.

Toma de decisiones: Esta conjunto de tareas se encarga de implementar las metodologías necesarias para la toma de decisiones, que conllevan a la mejora del contexto donde es aplicado el ciclo. Este conjunto de tareas, impactan la dinámica del ciclo con la intención de mejorarlo. Esto es evaluado nuevamente en las etapas de observación y análisis.

2.6 Salón de Clase Inteligente (SaCI)

Un Salón de Clase Inteligente puede ser definido como un entorno inteligente que guía un proceso de aprendizaje, considerando aspectos específicos del mismo (por ejemplo, perfil de aprendizaje de los estudiantes, curso impartándose, etc.). En [7,10,11] se ha desarrollado un middleware basado en sistemas multi-agentes, para permitir la autogestión de SaCI. Este middleware propone 3 niveles, uno para la administración de la comunidad de agentes, otro para administrar el acceso a servicios, aplicaciones, etc., y el último para caracterizar los diferentes componentes (software y hardware) de SaCI (Ver figura 2.10).

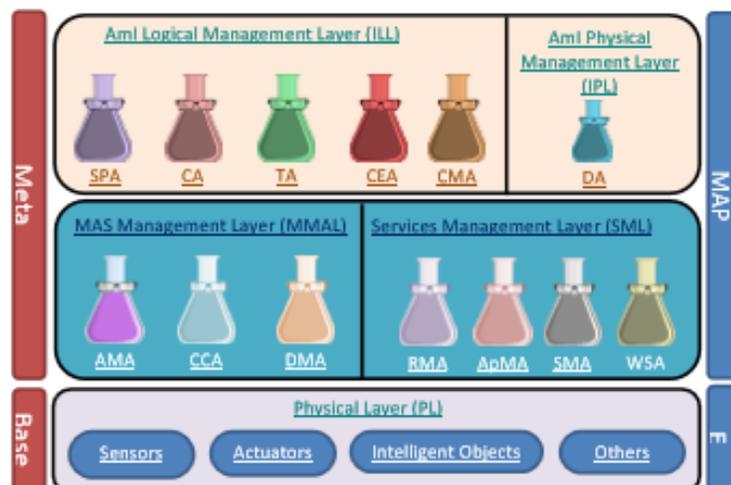


Figura 2.10: Middleware AmICL.

Es específico, a continuación se describen los componentes del middleware [2]:

Physical Layer (PL): es una capa física, con componentes que interconectan los elementos del ambiente inteligente (software o hardware), tales como APIs, librerías, etc. Este nivel trabaja directamente con el sistema operativo.

MAS Management Layer MMAL: es el nivel de gestión del SMA, que consiste en la comunidad de agentes de SaCI. Está compuesto por tres agentes (AMA, CCA y DMA), para la gestión de los diferentes aspectos requeridos por una comunidad de agentes (identificarse, comunicarse, etc.).

Service Management Layer SML: es el responsable de la búsqueda de servicios requeridos en la nube. Está formado por cuatro agentes, que permiten la comunicación bidireccional entre la comunidad de agentes y los servicios en la nube.

IE Logical Layer Management (ILL): especifica todas las aplicaciones de software e individuales presentes en SaCI como agentes, las cuales contienen metadatos que las definen. Entre los componentes principales de software se encuentran: VLE, sistema académico, sistema de recomendación de recursos educativos abiertos (REA), etc.; estos componentes proporcionan datos sobre el proceso de aprendizaje.

IE Physical Layer Management (PLL): especifica todos los dispositivos presentes en SaCI como agentes, las cuales contienen metadatos que las definen.

En general, SaCI es un aula inteligente centrada en el estudiante, que apoya el proceso de enseñanza-aprendizaje a través de dispositivos colaborativos y aplicaciones que facilitan la autoformación. Particularmente, en SaCI se genera una gran cantidad de datos sobre los procesos de aprendizaje, que pueden ser explotados por sus agentes para mejorar el proceso de aprendizaje. Por lo tanto, SaCI debe usar las tareas de LA para extraer el conocimiento oculto en los datos, a fin de producir información útil sobre los estudiantes, los recursos de aprendizaje, entre otras cosas, que pueden ser utilizados por los agentes de SaCI.

Capítulo 3

Especificación de los Ciclos Autonómicos

En este capítulo se describe el diseño conceptual de los ciclos autonómicos, además, se especifican las tareas que lo componen, los objetivos que cumplen y sus modelos de datos. El proceso fue realizado siguiendo la metodología MIDANO.

En cada ciclo autonómico se describen sus objetivos y se caracterizan sus tareas de AdD para optimizar algún aspecto del proceso que se da en SaCI. Las tareas de cada ciclo autonómico están agrupadas en Observación del proceso de aprendizaje, tareas que analizan el proceso de aprendizaje, y finalmente, tareas para tomar decisiones para mejorar el proceso de aprendizaje.

3.1 Ciclo 1: Determinar el paradigma de aprendizaje adecuado para un curso.

Objetivo:

- El objetivo de este ciclo es definir el paradigma de aprendizaje adecuado a ser utilizado en una clase durante un curso, basado en los datos obtenidos de los estudiantes.

Protocolo y Actividades: aquí se describe cada tarea de AdD de ese ciclo.

Especificación de las tareas de analítica de datos

1. Determinar cómo los estudiantes trabajan en las redes sociales.

2. Determinar cómo los estudiantes navegan en la web.
3. Determinar el rendimiento de los estudiantes.
4. Agrupar a los estudiantes por rendimiento, estilo de aprendizaje actual, y hábitos de internet.
5. Determinar el nuevo estilo de aprendizaje adecuado para un curso.

Las primeras tres tareas son tareas de observación, las cuales permiten monitorear el comportamiento de los estudiantes en diferentes contextos; la cuarta tarea es una tarea de análisis, hecha para interpretar el rendimiento, el comportamiento y hábitos de los estudiantes que conforman el grupo; finalmente, la última tarea es de toma de decisiones, la cual define si es necesario cambiar el estilo de aprendizaje actual o establecer uno nuevo para el curso o grupo, en base a los resultados obtenidos.

A continuación se describen las tareas y actividades del ciclo, la tabla 1 muestra las tareas, las fuentes de datos y las técnicas de analítica de datos que serán usadas en el ciclo.

Nombre de la Tarea	Descripción de la tarea		
	Fuente de datos	Técnica de analítica de datos	Rol
1. Determinar cómo los estudiantes trabajan en las redes sociales	Redes Sociales (Facebook, Twitter, etc.)	Técnicas de analítica social de datos (SNA)	Observación
2. Determinar cómo los estudiantes navegan en la web	Traza de navegación de internet	Minería Web	Observación
3. Determinar el rendimiento de los estudiantes	Base de datos del sistema académico	Minería de datos	Observación
4. Agrupar a los	Resultados	Minería de datos	Análisis

estudiantes por rendimiento, estilo de aprendizaje actual y hábitos de internet.	previos		
5. Determinar el Nuevo estilo de aprendizaje adecuado para un curso	Resultados previos	Minería de datos	Decisión

Tabla 3.1: Desarrollo de tareas de analítica de datos en el Ciclo 1.

El Ciclo autonómico se especifica en la Figura 3.1

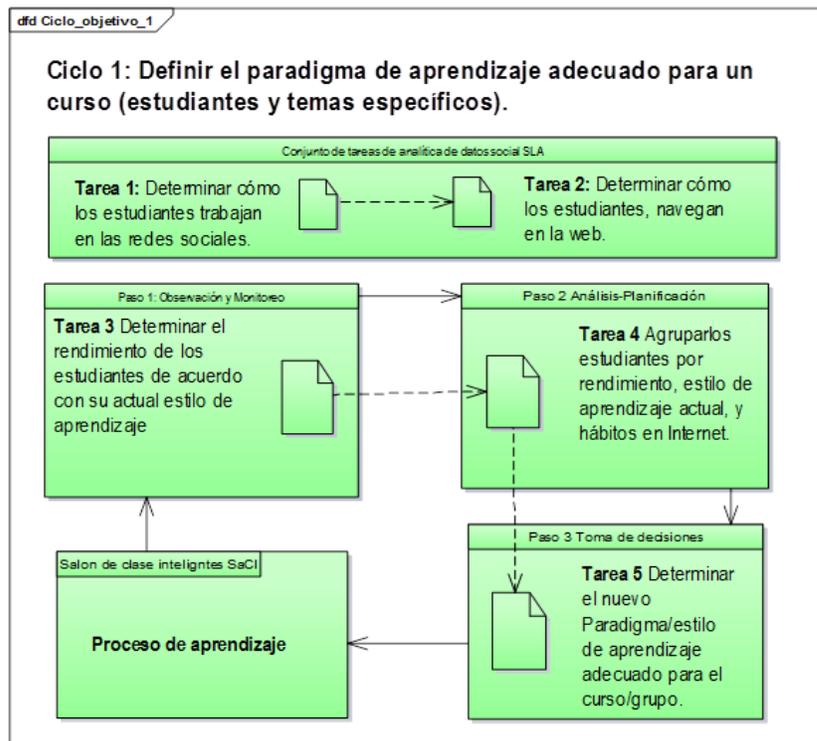


Figura 3.1: Esquema conceptual del Ciclo 1.

Paso 1: Observación

- Tarea 1: Determina como los estudiantes trabajan en las redes sociales: El objetivo específico de esta tarea es determinar el comportamiento de los estudiantes en las redes sociales que utilizan.
- Tarea 2: Determinar cómo los estudiantes navegan en la web. La idea principal detrás de esta tarea es descubrir el comportamiento de los estudiantes cuando navegan en la web, a través de las páginas web que visitan y la iteración que tienen en cada una de ellas, entre otras cosas.
- Tarea 3: Determinar el rendimiento de los estudiantes. En este caso se hará uso del sistema académico para identificar el puntaje obtenido por los estudiantes del curso en que estemos interesados, para luego relacionar el estilo de aprendizaje

Paso 2: Análisis

- Tarea 4: Agrupar estudiantes por rendimiento, estilo de aprendizaje y hábitos de internet. En esta tarea se lleva a cabo el agrupamiento de los estudiantes de acuerdo a tres descriptores, su estilo de aprendizaje actual, su rendimiento académico y sus hábitos en internet. Eso permite generar un patrón de grupo, para determinar el estilo de aprendizaje y los hábitos que deben ser considerados común a cada grupo, según su rendimiento académico.

Paso 3: Toma de decisiones

- Tarea 5: Determinar el nuevo estilo de aprendizaje adecuado para un curso. De acuerdo con los resultados previos, el mejor estilo es seleccionado según el grupo con el más alto rendimiento.

Definición del modelo de datos

El modelo de datos de este ciclo se aprecia en la figura 3.2. En ella se puede apreciar una vista parcial de la tabla de hechos con sus respectivas tablas de dimensión. Este modelo dimensional es donde serán almacenados los datos requeridos para cada tarea del ciclo autonómico.

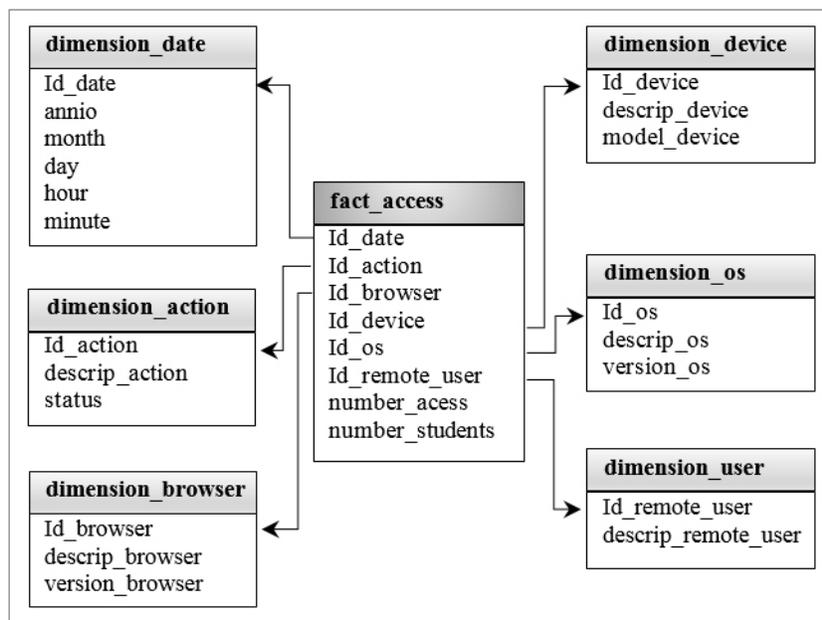


Figura 3.2: Modelo de datos del primer Ciclo Autonómico.

Este modelo de datos incluye los datos de la organización y los datos semánticos extraídos desde diferentes fuentes semánticas (Redes Sociales, Páginas Web, Linked Data, etc.). Cada tipo de información está incluida en una dimensión diferente en el modelo de datos de acuerdo con sus características, por ejemplo:

dimension_user: se almacena la información de los estudiantes del curso, su record académico, asignatura y preferencias en la plataforma, entre otras cosas. Puede ser extendida para incorporar información requerida para el logro de otro objetivo de otro ciclo.

dimension_browser: aquí es almacenada la información del comportamiento de los estudiantes en relación a los navegadores web, y es extendida para almacenar el

comportamiento semántico relacionado a la minería de uso, minería web y minería de texto, entre otras.

dimension_device: en esta dimensión se almacena los dispositivos desde los cuales los usuarios accesan a la plataforma computacional, puede ser extendida igual que las otras dimensiones para obtener el conocimiento semántico de interés.

dimensión_os: en esta dimensión se almacenan los sistemas operativos desde los cuales los estudiantes trabajan.

dimension_date: donde se almacena la información relacionadas a las fechas y horas en las que los usuarios realizan sus actividades. Esta dimensión proporciona conocimiento semántico relacionado a preferencias de horas, días de los usuarios.

dimension_action: almacena los datos relacionados a las acciones realizadas por los usuarios en el sistema, proporciona información semántica del comportamiento de los usuarios.

3.2 Ciclo 2: Determinar el recurso educativo ideal para un estudiante

Objetivo:

- El objetivo de este ciclo es determinar el recurso educativo ideal para un estudiante en un momento dado.

Protocolo y Actividades: las tareas de analítica de datos son:

1. Clasificar a los estudiantes por su rendimiento y participación.
2. Determinar la relación entre la clasificación de los estudiantes y el recurso educacional utilizado.
3. Determinar el mejor recurso educativo para un estudiante, de acuerdo a los resultados previos y el estilo de aprendizaje utilizado.

La primera es una tarea de observación la cual determina el rendimiento de los estudiantes, la segunda es una tarea de análisis que define el mejor recurso educacional según los estudiantes que las usan, y finalmente, la tarea de toma de decisiones define el recurso educacional adecuado para un estudiante. La estructura de este ciclo se muestra en la figura 3.3

A continuación se describen las tareas necesarias para este ciclo. La tabla 2 muestra las tareas, las Fuentes de datos y las técnicas de analítica de datos que serán utilizadas en el ciclo 2.

Nombre de la tarea	Descripción de la tarea		
	Fuente de datos	Técnica de analítica de datos	Rol
1. Clasificar a los estudiantes por sus calificaciones y participación	VLE, Sistema académico, etc.	Minería de Datos	Observación
2. Determinar la relación entre la clasificación de los estudiantes y el recurso educacional utilizado	Resultados previos, VLE, repositorios de recursos educativos	Minería de Datos	análisis
3. Determinar el mejor recurso educacional para un estudiante dado de acuerdo a los resultados previos y su estilo de aprendizaje	Resultados previos, sistema académico y repositorios de recursos educativos	Minería de Datos	Decisión

Tabla 3.2: Descripción de tareas en el Ciclo 2.

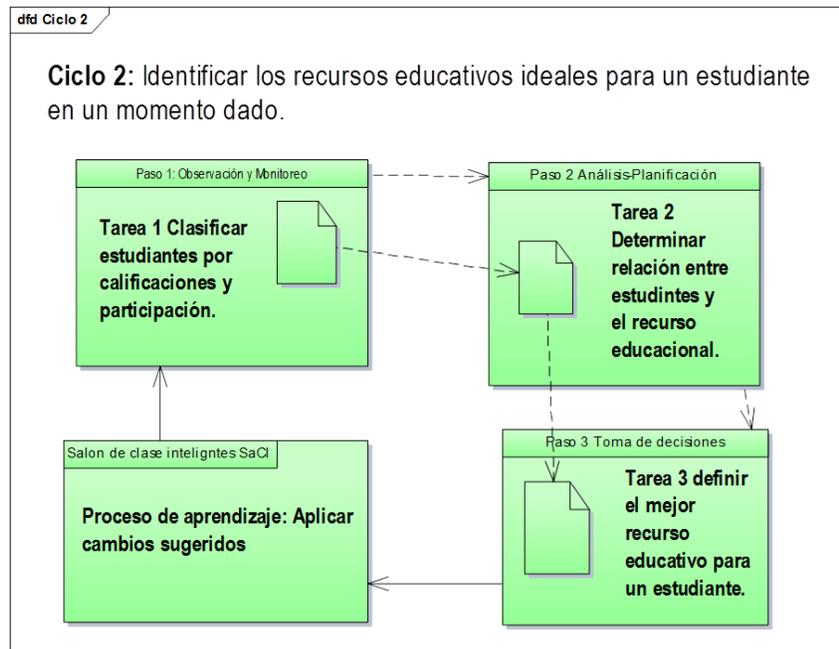


Figura 3.3: Esquema conceptual de Ciclo 2.

Paso 1: Observación

- Tarea 1: Clasificar a los estudiantes por sus calificaciones y participación. En esta tarea, el objetivo es establecer un ranking entre los estudiantes de acuerdo a su rendimiento.

Paso 2: Análisis

- Tarea 2: Determinar la relación entre la clasificación de los estudiantes y el recurso educacional utilizado. En esta tarea, el objetivo es establecer la relación entre buenos rendimientos y el recurso educacional utilizado.

Paso 3: Toma de decisiones

- Tarea 3: Determinar los recursos educacionales utilizados por los mejores estudiantes, para un estilo de aprendizaje dado. Esta tarea toma la decisión de cuál de todos los recursos educacionales que se han

utilizado es el mejor para un estudiante dado, según su estilo de aprendizaje. La decisión es tomada evaluando diferentes aspectos: las características del estudiante, el recurso educacional usado por los estudiantes con mejores rendimientos académicos según los estilos de aprendizaje, entre otras cosas.

Definición del modelo de datos

En este caso, y para los siguientes ciclos, el modelo de datos del primer ciclo puede ser usado para obtener el objetivo, para ello es necesario agregar a la tabla de hechos las métricas generadas en cada ciclo, por ejemplo: la clasificación de los estudiantes, los mejores recursos educativos por estilo de aprendizaje, entre otros.

3.3 Ciclo 3: Identificar a los estudiantes con necesidades específicas

Objetivo:

- El objetivo de este ciclo es definir cuáles son los estudiantes que requieren atención particular durante el proceso de aprendizaje. Este ciclo autónomo está compuesto por las tareas siguientes (ver figura 4).

Protocolo y Actividades: Especificación de las tareas de analítica de datos

1. Clasificar los estudiantes por rendimiento académico.
2. Determinar las tareas/actividades en donde los estudiantes han tenido éxito y las que no.
3. Determinar en donde los estudiantes fallaron.
4. Buscar un recurso educativo para ayudar a los estudiantes en las tareas/actividades que fallaron.

5. Poner juntos esos recursos educacionales obtenidos en el paso anterior de acuerdo a una secuencia lógica.

La primera tarea es similar al ciclo autonómico anterior, la segunda y la tercera analizan los problemas donde los estudiantes fallaron, y la última prepara un nuevo recurso educacional para los estudiantes con necesidades especiales, y así ayudarlos a mejorar el rendimiento académico.

A continuación se describen las tareas necesarias para este ciclo. La tabla 3.3 muestra las tareas, las Fuentes de datos, y las técnicas de analítica de datos que serán utilizadas en el ciclo 3.

Nombre de la tarea	Descripción de la tarea		
	Fuente de datos	Técnica de analítica de datos	Rol
Clasificar los estudiantes por rendimiento académico	VLE, Sistema académico, etc.	Minería de datos	Observación
Determinar las tareas/actividades en donde los estudiantes han tenido éxito y las que no	Resultados previos, VLE, repositorios de recursos educacionales	Minería de datos	análisis
Determinar en donde los estudiantes fallaron	Resultados previos, VLE, repositorios de recursos educacionales	Minería de datos	análisis
Buscar un Nuevo recurso educativo para ayudar a los estudiantes en las tareas/actividades que fallaron	Resultados previos, internet	Linked Data	Decisión

Poner juntos esos recursos educativos obtenidos en el paso anterior de acuerdo a una secuencia lógica	Resultados previos.	Minería de datos	Decisión
--	---------------------	------------------	----------

Tabla 3.3. Descripción de las tareas del Ciclo 3.

Producto:

Esquema conceptual del ciclo autónomo:

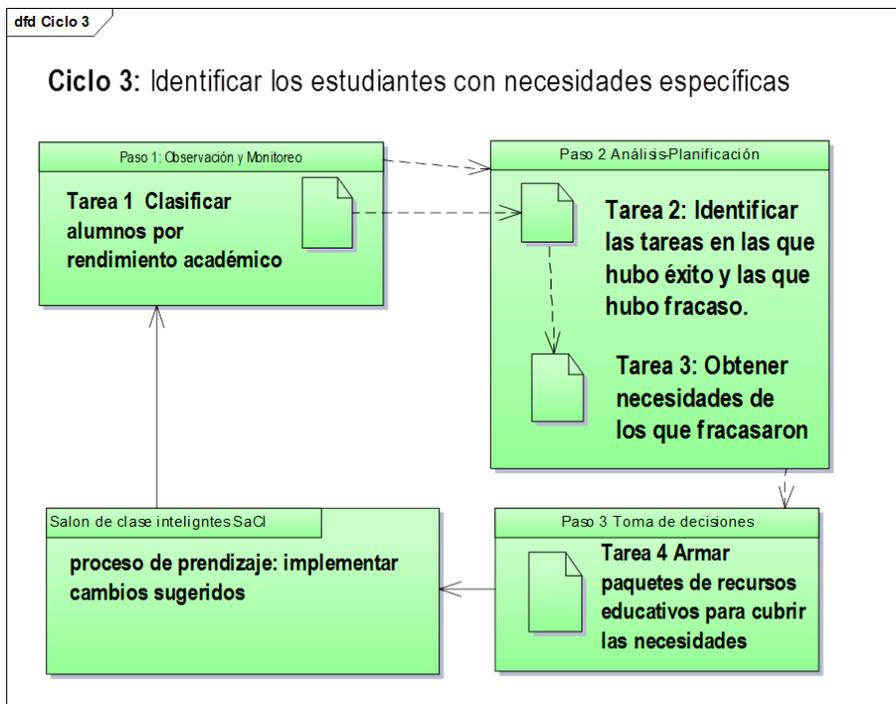


Figure 3.4: Esquema conceptual del Ciclo 3.

Paso 1: Observación

- Tarea 1: Clasificar los estudiantes por rendimiento académico. El objetivo a lograr de esta tarea es clasificar a los estudiantes por rendimiento académico.

Paso 2: Análisis

- Tarea 2: Determinar las tareas/actividades en donde los estudiantes han tenido éxito y las que no. El objetivo de esta tarea es a identificar las actividades donde los estudiantes con necesidades especiales obtuvieron resultados favorables y en las que no. La idea general es identificar las actividades en las que los estudiantes lograron mejores resultados y las actividades en las que los estudiantes fallaron, para xxx.
- Tarea 3: Determinar necesidades de los estudiantes: El objetivo de esta tarea es identificar las necesidades específicas de los estudiantes, teniendo en cuenta en donde han fallado. Esta tarea usa lo resultado de la tarea anterior y determina los requerimientos, las características y actividades donde os estudiantes fallaron para establecer sus necesidades.
- Tarea 4: Buscar un recurso educativo para ayudar a los estudiantes en las tareas/actividades que fallaron. En esta tarea se usan técnicas para buscar recursos educativos en la web, las cuales pueden ayudar a los estudiantes a cubrir sus necesidades.

Paso 3: Toma de decisiones.

- Tarea 5: Poner juntos los recursos educacionales identificados en el paso anterior, de acuerdo a una secuencia lógica..

Modelo de datos

El modelo de datos de este ciclo es el mismo usado en el ciclo anterior.

3.4 Ciclo 4: Evitar la deserción estudiantil

Objetivo:

- El objetivo de este ciclo es prevenir la deserción estudiantil, analizando el rendimiento y comportamiento de los estudiantes, para predecir el grupo de

potenciales desertores con suficiente tiempo, para planificar una estrategia para evitarlo.

Protocolo y Actividades: Especificación de las tareas de analítica de datos

1. Clasificar los estudiantes en desertores o no desertores.
2. Predecir el desertor potencial.
3. Crear un patrón de deserción.
4. Evitar la deserción estudiantil.

La primera es una tarea de observación, la cual identifica los potenciales desertores de un curso; con esta información, la segunda tarea construye un modelo predictivo de los estudiantes que pueden desertar. La siguiente tarea, de análisis, construye el patrón de estos desertores potenciales, y finalmente, la última tarea, de toma de decisiones, define las estrategias que deben ser aplicadas para evitar esta deserción de los estudiantes.

A continuación se describen las tareas de este ciclo. La tabla 3.4 muestra las tareas, Fuentes de datos y técnicas de analíticas de datos que serán usadas para completar este ciclo.

Nombre de la tarea	Descripción de las tareas		
	Fuente de datos	Técnica de analítica de datos	Rol
1. Clasificar a los estudiantes en desertores o no desertores	Sistema Académico	Minería de datos	Observación
2. Predecir el desertor potencial	Resultados previos	Minería de datos	Análisis
3. Crear un patrón de deserción	Resultados previos	Minería de datos	Análisis
4. Evitar la deserción estudiantil	Resultados previos, Internet.	Minería de datos, Minería web	Decisión

		Minería de texto	
--	--	------------------	--

Tabla 3.4. Descripción de las tareas del Ciclo 4.

Producto:

Esquema conceptual del ciclo autonómico:

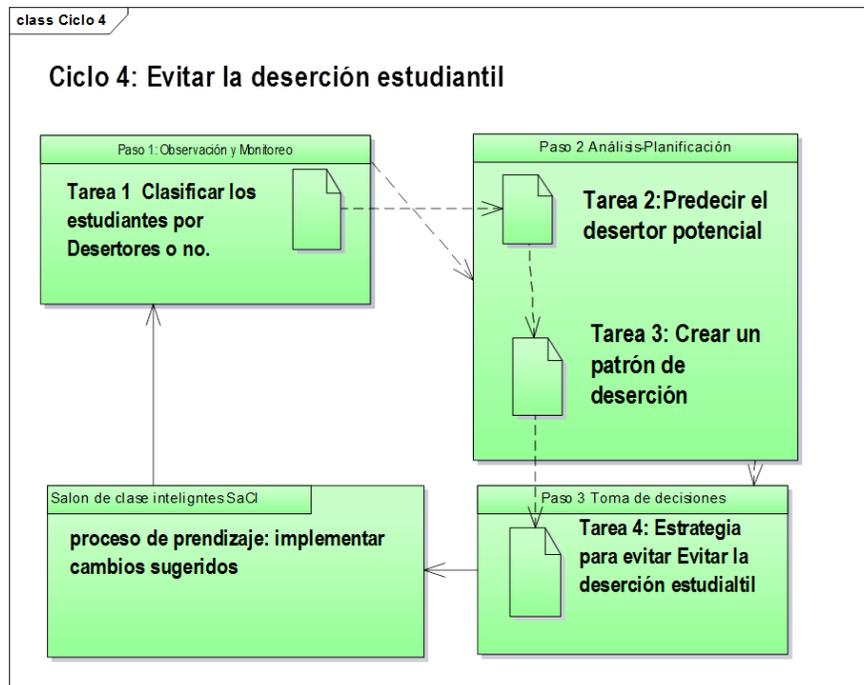


Figura 3.5: Esquema conceptual del Ciclo 4.

Paso 1: Observación

- Tarea 1: Clasificar a los estudiantes por desertores o no desertores. El objetivo de esta tarea es clasificar a los estudiantes, de acuerdo a si estos son desertores o no.

Paso 2: Análisis.

- Tarea 2: Predecir el desertor potencial. El objetivo de esta tarea es determinar los potenciales desertores. Para este trabajo la tarea construye un modelo predictivo usando información tomada de la tarea anterior.
- Tarea 3: Crear un patrón de deserción: El objetivo de esta tarea es encontrar los patrones que caracterizan a los desertores. En este sentido, define el perfil de cada grupo de desertores. Esta tarea usa los resultados de la tarea anterior para determinar los patrones.

Paso 3: Toma de decisiones.

- Tarea 4: Evitar la deserción estudiantil. En esta tarea son seleccionadas las mejores técnicas y actividades de enseñanza que pueden prevenir la deserción estudiantil, estas técnicas se pueden inferir a partir de los patrones de deserción obtenidos en la tarea anterior.

Capítulo 4

Desarrollo del Ciclo Autonomico “Definir el paradigma de aprendizaje adecuado para un Curso”

En este capítulo es especificado uno de los ciclos autonomico de tareas de analítica de datos de SaCI, para evaluar su factibilidad. Se ha seleccionado el ciclo “Definir el paradigma de aprendizaje adecuado para un Curso”, debido a que posee tareas de analítica de datos social y de analítica de datos organizacional.

4.1 Caracterización General del Ciclo Autonomico

Objetivo General del “Ciclo Autonomico”

Definir el paradigma de aprendizaje adecuado para un curso en SaCI.

Objetivos específicos:

- Utilizar Analítica Social de Aprendizaje (SNA) para determinar el estilo de aprendizaje adecuado para los estudiantes del curso seleccionado.
- Determinar el rendimiento académico de los estudiantes del curso, para determinar los estilos de aprendizaje con mejores tasas de éxito.
- Determinar la lista de actividades ofrecidas por una materia en un curso.
- Determinar la lista de actividades utilizadas por los estudiantes en las redes sociales.
- Determinar la lista de actividades ideal para los estudiantes de un curso determinado.

4.2 Especificación del Ciclo Autonomico

Como puede apreciarse en el capítulo 3, los ciclos se especifican en un proceso secuencial. A continuación es presentado este ciclo con detalle:

4.2.1 Representación Gráfica del Ciclo Autonomico.

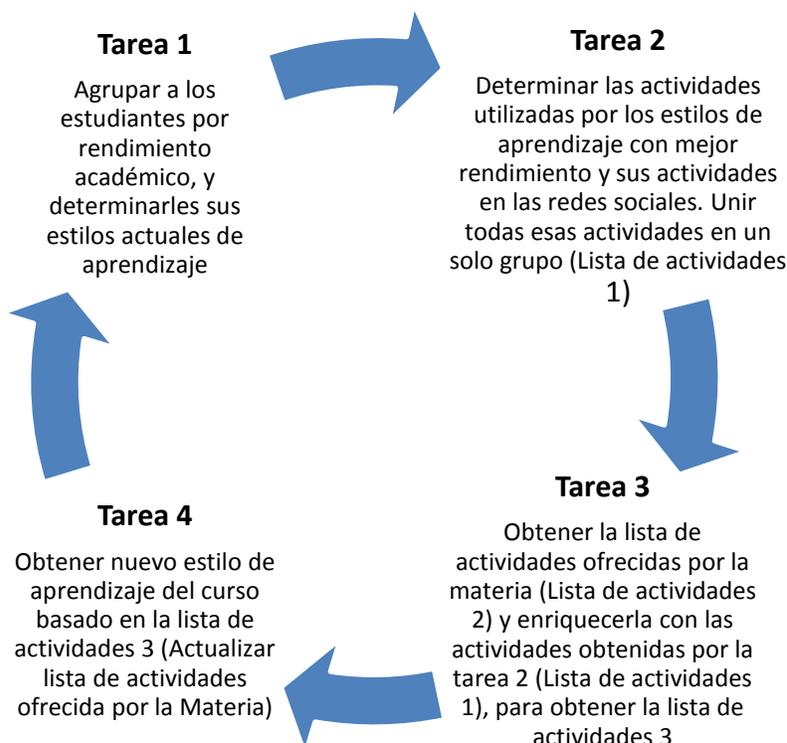


Figura 4.1: representación gráfica del Ciclo Autonomico: “Determinar el paradigma de aprendizaje adecuado para un curso”.

4.2.2 Tareas que componen el Ciclo Autonomico.

- Tarea 1: Agrupar a los estudiantes por rendimiento académico, y determinarles sus estilos actuales de aprendizaje
- Tarea 2: Determinar las actividades utilizadas por los estilos de aprendizaje con mejor rendimiento, y sus actividades en las redes sociales. Unir todas esas actividades en un solo grupo (Lista de actividades 1)

- Tarea 3. Obtener la lista de actividades ofrecidas por la materia (Lista de actividades 2) y enriquecerla con las actividades obtenidas por la tarea 2 (Lista de actividades 1), para obtener la lista de actividades 3
- Tarea 4: Obtener nuevo estilo de aprendizaje del curso basado en la lista de actividades 3 (Actualizar lista de actividades ofrecida por la Materia)

4.2.3 Descripción de las tareas del ciclo autónomo actual

Descripción de la Tarea 1

Nombre de la tarea	Agrupar a los estudiantes por rendimiento académico, y determinarles sus estilos de aprendizaje actual
Descripción	Calcular el rendimiento de los estudiantes, y buscar en las BD de SaCI cuál es su actual estilo de aprendizaje (ya que se puede ir actualizando en el tiempo).
Fuentes de datos	Base de Datos de SaCI,
Tipo de técnica utilizar	Analítica de datos clásica.
Técnica de analítica de datos	Agrupación
Tipo de tarea de analítica de datos	
Tipo de modelo	Modelo de Agrupamiento K-Means. Ver apéndice H
Métricas	
Tareas relacionadas con la analítica de datos	-Obtener record académico de los estudiantes. -Agrupara los estudiantes por rendimiento. -Obtener los grupos de mejor rendimiento.
Insumo generado y tarea que lo consume	Insumo generado: Estudiantes con su Rendimiento académico y su patrón de estilo de aprendizaje ordenado. Tarea que lo consume: (Tarea 2)
Tipo de tarea del ciclo autónomo (rol)	Observación.

Tabla 4.1: Descripción de la tarea 1.

Descripción de la Tarea 2

Nombre de la tarea	Determinar las actividades académicas utilizadas por
---------------------------	--

	los estilos que obtuvieron mejor rendimiento y sus actividades en las redes sociales. Unir todas esas actividades en un solo grupo.
Descripción	Determinar la lista de actividades de SLA realizadas de preferencia por los estudiantes de alto rendimiento en las redes sociales. Determinar las actividades usadas según sus estilos de aprendizaje por los estudiantes de alto rendimiento. Unirlas en una sola lista de actividades.
Fuentes de datos	Twitter, Facebook, etc. y BD de SaCI con las actividades realizadas en el curso por los estudiantes.
Tipo de técnica utilizar	(SNA) para determinar con minería de texto y Minería de uso de la Web, la lista de actividades en las redes sociales.
Técnica de analítica de datos	Minería de texto y de uso de la Web
Tipo de tarea de analítica de datos	Búsqueda de Patrones.
Tipo de modelo	Modelo Descriptivo.
Métricas	Frecuencias de las actividades web.
Tareas relacionadas con la analítica de datos	-Minería de texto –Frecuencia de palabras clave (Actividades). - Minería de uso de la Web - Frecuencia de comportamientos
Insumo generado y tarea que lo consume	Insumo generado: Lista de actividades 1. Tarea que lo consume: (Tarea 3).
Tipo de tarea del ciclo autónomico (rol)	Observación.

Tabla 4.2: Descripción de la tarea 2.

Descripción de la Tarea 3

Nombre de la tarea	Obtener la lista de actividades ofrecidas por la materia (Lista de actividades 2) y enriquecerla con las actividades obtenidas por la tarea 2 (Lista de actividades 1), para obtener la lista de actividades 3.
Descripción	Esta tarea se ocupa de obtener la lista de actividades ofrecidas por la materia que será la Lista de actividades 2, y compararla con las actividades

	obtenidas en el paso anterior (Lista de actividades 1), para obtener la lista de actividades 3.
Fuentes de datos	Base de Datos de SaCI
Insumo generado y tarea que lo consume	Insumo generado: Lista de actividades necesarias para el curso. Tarea que lo consume: (Tarea 4).
Tipo de tarea del ciclo autónomico (rol)	Análisis.

Tabla 4.3: Descripción de la tarea 3.

Descripción de la Tarea 4

Nombre de la tarea	Obtener nuevo estilo de aprendizaje del curso basado en la lista de actividades 3
Descripción	Esta tarea se ocupa de actualizar el estilo de aprendizaje adecuado para la materia.
Fuentes de datos	Tarea 4
Insumo generado y tarea que lo consume	Insumo generado: Estilo de aprendizaje adecuado para el curso. Tarea que lo consume: los otros componentes de SaCI, como el agente recomendador de Objetos de Aprendizaje, entre otros
Tipo de tarea del ciclo autónomico (rol)	Decisión

Tabla 4.4: Descripción de la tarea 4.

4.2.4 Macro-Algoritmo del Ciclo Autónomico de Tareas

Macro Algoritmo del ciclo autónomico

Tarea	Paso	Detalle	Tarea de Minería	Herramientas	Producto
1	P1	Obtención de los datos de los estudiantes SaCI	-	SQL	Archivo .csv
1	P2	Agrupar por rendimiento académico.	Agrupación	Weka o R	Estudiantes con su Rendimiento académico y patrón de estilo de

					aprendizaje(Si lo hay) Archivo .arff o Archivo csv
2	P3	Obtención de lista de actividades académicas de los mejores rendimientos		SQL	Lista de actividades actuales de los estudiantes
2	P4	Obtención de la lista de actividades realizadas en las redes sociales	Minería de texto en redes sociales (Twitter, etc.) y de uso en la Web	R	Lista de actividades Sociales
2	P5	Unir las dos listas de actividades			Lista de actividades 1
3	P6	Obtención de lista de actividades ofrecidas por la materia		SQL	Lista de actividades 2
3	P7	Realizar la mezcla entre las actividades de las lista 1 y 2			Lista de actividades 3
4	P8	Determinar la tabla de frecuencias de las actividades de cada dimensión del modelo de modelo Felder Silverman ver apéndice F.	Calcular el número de actividades realizadas en cada dimensión del modelo Felder Silverman, contenidas en la tabla de frecuencias	R (Ver Apéndice D)	Tabla General de Frecuencias de actividades
4	P9	Determinar el estilo de aprendizaje del curso a partir de la lista de actividades 3		R	Estilo de aprendizaje

Tabla 4.5: Macro – Algoritmo del Ciclo Autónomo.

4.3 Incorporación del CA en el PTL de SaCI.

La Figura 4.2 representa el modelo de conversación del Proceso de Tutoría en Línea (PTL) de SaCI, donde podemos apreciar la iteración entre los agentes de SaCI (el tutor, el alumno, el VLE, el CA, el SR, etc.) que participan en este proceso. En la Figura 4.2 se observa que la primera actividad es la inicialización de la sesión por parte del (de los) estudiante(s). Esa sesión es verificada por el VLE, que invoca el servicio de “obtener el Perfil del Estudiante (PE) al Ciclo Autónomo (CA). Este PE lo calcula el CA y se lo envía como respuesta de la invocación a ese servicio al VLE, el cual después lo usa para establecer las actividades a realizar durante el curso y se lo envía al Sistema Recomendador (RS). El RS, basado en el PE, realiza la búsqueda en Internet de los objetos de aprendizaje adecuados al PE para ese curso/estudiante. Esta información que encuentra el RS se la envía al VLE (la información sobre los Recursos Educativos de Aprendizaje (LER) recomendados), que planifica la ejecución de los LERs en el marco del PTL. El tutor supervisa dicha planificación, y sujeto a la aprobación de él, el VLE lo envía al Smart Board (SB). El SB despliega los LERs según el plan, en un ciclo repetitivo de interacción entre el estudiante, el SB y el tutor. Al final de la sesión del curso, es actualizado el PE de los estudiantes y/o del curso, y el VLE le notifica al Sistema Académico actualizar las calificaciones de los estudiantes, de ser necesario.

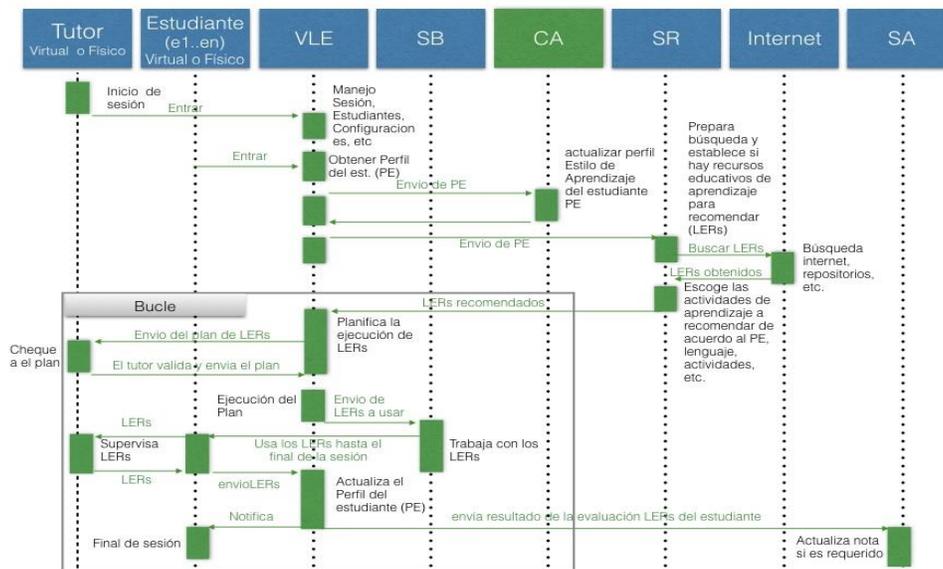


Figura 4.2: Conversación del PTL SaCI usando el servicio provisto por el CA.

Capítulo 5

Escenarios Experimentales y Análisis de Resultados para el Ciclo Autonomico “Definir el paradigma de aprendizaje adecuado para un Curso”

En este capítulo se experimenta con el ciclo autonómico de tareas implementado en el capítulo anterior, en diferentes escenarios o contextos, para después analizar los resultados obtenidos. Para todos los escenarios, la información usada proviene de la base de datos de Moodle (Ver apéndice G) de la Universidad Técnica Particular de Loja UTPL del Ecuador, a quien le agradecemos por haber posibilitado el uso de la misma.

5.1 Escenarios Experimentales para el Ciclo Autonomico

En esta sección se definen varios escenarios experimentales, con el fin de analizar el comportamiento del ciclo autonómico “descubrir paradigma de aprendizaje adecuado de un Curso en SaCI”, descrito en el paso anterior. A continuación se definen los escenarios utilizados:

5.1.1 Escenario de prueba Nro. 1.

Objetivo a medir

Determinar el estilo de aprendizaje de un curso de Computación en SaCI, partiendo del hecho de que es desconocido.

Características del contexto

Se trata de un curso de computación, en el que los estudiantes están registrados en la plataforma Moodle con sus Id, sus respectivas notas finales de los cursos anteriores, entre otras cosas, pero sin conocerse sus actuales estilos de aprendizaje (caso común, al inicio en las bases de datos de Moodle). En este caso, queremos determinar el estilo de aprendizaje de los mejores estudiantes, a partir de sus notas previas, las actividades que realizan, y sus comportamientos en las redes sociales. Esta tarea es fundamental para SaCI, porque le permitirá establecer las herramientas de aprendizaje adecuadas, que deberá proveer a los estudiantes del curso de computación, a partir del perfilamiento de los estilos de aprendizaje de los mejores estudiantes. Esto constituye una especie de consenso social sobre el estilo de aprendizaje que deberá seguir SaCI. También son contemplados los estudiantes con requerimientos especiales, para ellos, se tienen otros ciclos autónomos que los ayudarían en sus procesos de aprendizaje.

Aplicación del Macro-Algoritmo

La tabla 5.1 representa la instanciación del macro-algoritmo del ciclo autónomo, presentado en la sección 4.2.4, en el caso de estudio.

Tarea	Paso	Detalle	Tarea de Minería de Datos	Herramientas usada	Producto
1	E1P1	Obtención de los datos de los estudiantes SaCI	-	SQL	Lista de Estudiantes
1	E1P2	Agrupar estudiantes por rendimiento académico.	Agrupamiento de los estudiantes según sus rendimiento histórico	SPSS Agrupamiento basado en la distancia K-Means.	Estudiantes agrupados en: Alto Rendimiento Bajo rendimiento Medio Rendimiento
2	E1P3	Obtención de lista de actividades académicas realizadas por los		SQL	Lista de actividades académicas actuales de los estudiantes de alto rendimiento.

		estudiantes de mejores rendimientos, basados en el modelo Felder Silverman.			
2	E1P4	Obtención de la lista de actividades realizadas en las redes sociales, por los estudiantes, basada en el modelo Felder Silverman	Minería de texto en Twitter. Filtro de actividades según estilos de aprendizaje de los estudiantes (ver apéndice E)	R (Ver Apéndice C)	Actividades y Frecuencia de las mismas
2	E1P5	Unir las dos listas de actividades			Lista de actividades académicas y Sociales de los estudiantes
3	E1P6	Obtención de lista de actividades ofrecidas por la materia		SQL	Lista de las actividades que en ese curso de computación se deben realizar.
3	E1P7	Realizar la mezcla entre las actividades de las listas anteriores			Lista de todas las actividades identificadas en los pasos anteriores
4	E1P8	Determinar la tabla de frecuencias de las actividades de cada dimensión del modelo de modelo Felder Silverman	Calcular el número de actividades realizadas en cada dimensión del modelo Felder Silverman, contenidas en la tabla de frecuencias	R (ver Apéndice D)	Tabla General de Frecuencias de actividades
	E1P9	Determinar el estilo de aprendizaje de la categoría del curso a partir de la lista de frecuencias de actividades general	Uso de la tabla de frecuencias para obtener el estilo de aprendizaje	R (ver Apéndice D)	Gráfico Estilos de Aprendizaje detectados

Tabla 5.1: Instanciación del Macro-Algoritmo para el escenario 1

E1P1: Obtención de datos de los estudiantes SaCI

La primera consulta consistió en conseguir las notas de los estudiantes que están actualmente en SaCI (ver figura 5.1).

```

select grades.userid, cats.id as materia, round(finalgrade)
from moodle19.mdl_enrol_utpl enrol
left join moodle19.mdl_course course on enrol.courseid = course.id
left join moodle19.mdl_course_categories cats on course.category =
cats.id
left join moodle19.mdl_grade_grades grades on enrol.userid =
grades.userid and FINALGRADE order by userid
    
```

Figura 5.1: Consulta a la base de datos para recuperar a los estudiantes que están en SaCI (que corresponden al curso de Computación, identificado como categoría 5284 en la base de datos).

La tabla 5.2 muestra las tablas de la base de datos involucradas en la consulta, y la descripción de los campos utilizados en ella.

Tablas Involucradas	Descripción
Moodle19.mdl_grade_grades	Contiene las notas mínimas, máximas y actuales asociadas a ítems. Un ítem puede ser un quiz, examen, asignación, etc.
Moodle19.mdl_enrol_utpl	Contiene los estudiantes inscritos en cursos y asignaturas de utpl
Moodle19.mdl_course_categories	Contiene la información de los cursos ofrecidos.
Moodle19.mdl_course_categories	Contiene información de las categorías o carreras a las que corresponde cada curso
Moodle19_mdl_assignment_submission	Contiene la información necesaria para determinar si una asignación fue entregada o no.

Campos Empleados	Descripción
Mdl_grade_grades.userid	Identificador de usuario asociado a una nota
Mdl_grade_grades.finalgrade	Valor obtenido por un alumno en un ítem
Mdl_grade_grades.rawgrademax	Valor máximo que podía obtenerse en un ítem
Mdl_enrol_utpl.courseid	Identificador de un curso asociado a un estudiante
Mdl_enrol_utpl.userid	Identificador de usuario inscrito en la plataforma

Tabla 5.2: Tablas y campos utilizados en las consultas SQL.

EIP2: Agrupar estudiantes por rendimiento.

Con la consulta obtenida en el paso anterior, se procede a agrupar a los estudiantes. El proceso fue modelado con el agrupamiento basado en distancias, específicamente el K-Means, para $K=3$ clusters. Así, la muestra fue dividida en tres cluster, aquellos que obtuvieron el mejor rendimiento en la materia, los cuales son los que corresponden al cluster 1 (Ver figura 5.2), con puntaje entre 80 y 100 en la nota final, los que obtuvieron puntaje entre 60 y 80, correspondientes al cluster 2, y el resto, al cluster 3. El agrupamiento fue realizado según el siguiente procedimiento (en el apéndice B está el programa en java que lo realiza):

1. Leer los datos del archivo csv.
2. Establecer la distancia entre el rendimiento académico de cada estudiante y el rendimiento de los estudiantes de cada grupo “alto rendimiento”, “medio rendimiento” y “bajo rendimiento”, y según eso, enviarlos a alguno de los grupos a los que pertenece.
3. Crear los archivos de salida con los rendimientos.

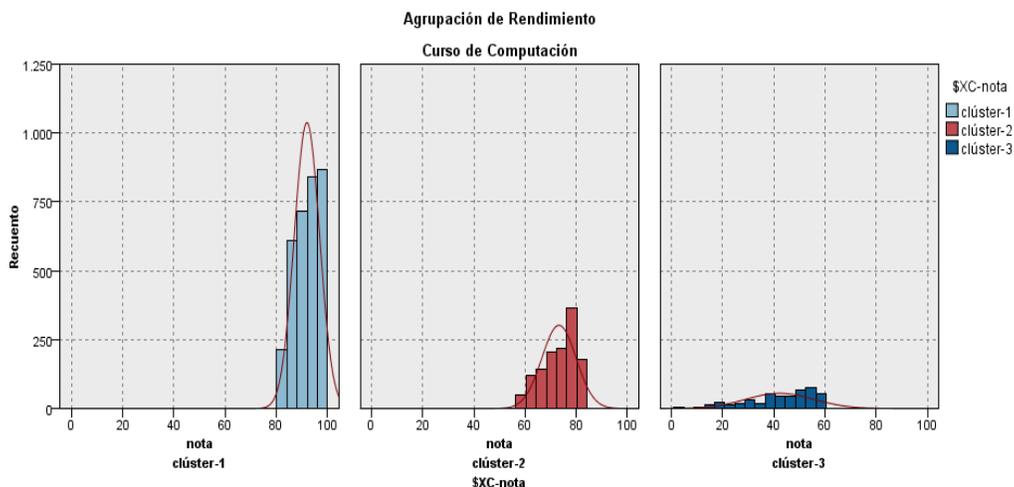


Figura 5.2: Agrupamiento K-Means del curso de computación.

El resultado del agrupamiento genera dos archivos:

- **estCursoClase.csv:** contiene todos los datos de la consulta SQL, y adicionalmente, contiene una columna que muestra donde quedó agrupado el estudiante.
- **estCursoClaseSi.csv:** contiene solo los datos de los estudiantes de alto rendimiento, a los que se les buscará la lista de actividades utilizadas durante el curso.

E1P3: Obtención de lista de actividades académicas realizadas por los estudiantes en SaCI.

A continuación es realizada la consulta en Moodle, de las actividades realizadas por los estudiantes de alto rendimiento en el curso de computación. Esas actividades académicas corresponden a estrategias, instrumentos educativos y herramientas de aprendizaje, del modelo de estilos de aprendizaje de Felder Silverman. El resultado es almacenado en 2 archivos, denominados: actividadesEstMoodle.csv y actividadesEstMoodle.txt.

E1P4: Obtención de la lista de actividades realizadas por los estudiantes en las redes sociales

En este paso del algoritmo, se utilizó el API de twitter, para extraer desde allí las actividades en las redes sociales por parte de los estudiantes. El resultado de este paso es una lista de archivos identificados con las cuentas de cada estudiante, uno txt, que contiene los últimos 100 tweets del estudiante, y otro csv que contiene las frecuencias de las actividades posteadas en la red social. Finalmente, se construye uno que une a todos, denominado todosRD.txt. El procedimiento de esta fase se explica a continuación:

- **Obtención de credenciales twitter:** para poder obtener acceso a los datos de twitter, es necesario registrarse en esta página como desarrollador, luego se usa `api_key`, `api_secret`, `access_token` y `access_token_secret`, para obtener los datos desde la red social.
- **Proceso de captura de tweets:** para este paso se escribió un Script en R (ver figura 5.3), que permite extraer datos relacionados con los usuarios, como por ejemplo: su estado, la lista de seguidores, el número de veces que son retwiteados sus mensajes, el número de tweets hechos por el usuario, entre otras cosas. En particular, en este trabajo seleccionamos la función `timeline()`, que devuelve la lista de tweets realizados por el usuario en la red (ver Figura 5.4).

```

1  api_key<-"x1XqNv2WUy8bicPU9oH9AYCE1"
2  api_secret<-"ZpZoEGmAp1qK1BnI9QM4QD5ZMLuNuLxctjsKAXHjFFGjLpV91v"
3  acces_token<-"163738165-02bwe1AZ4NYRuLOSVLeBeWE6RjlyamXPsv1y8hw8"
4  acces_token_secret<-"pjLWou5uBDTy15uj7NyVozm07FcJ7YtbvRnTsmA0qQoyN"
5  #Cargar libreria
6  library(twitteR)
7  #solicito autenticacion a twitter
8  setup_twitter_oauth(api_key,api_secret,acces_token,acces_token_secret)
9  # Solicitar primeros 100 tweets del timeline del estudiante
10 t <- userTimeline("estudianteUTPL", n = 100)
11 t<-twListToDF(t)
12 # obtiene el texto de los tweets
13 txt = t$text
14 txt2=t$screenName
15 result<-cbind(txt2,txt)
16 str(result)
17 write.csv(result,file="estudiantes.csv",row.names = F)
18

```

Figura 5.3: Script para la obtención del timeLine de los estudiantes

	text	favorited	favoriteCount	replyToSN	created	truncated	replyToSID	id
1	Foro. No olviden dejar los comentarios y las inquietud...	FALSE	0	NA	2018-02-14 19:36:45	FALSE	NA	9638595754547978...
2	Exposiciones. Los temas de las exposiciones est<U+...	FALSE	0	NA	2018-02-14 19:31:45	FALSE	NA	9638583191181598...
3	Material en PDF. El material correspondiente a las clas...	FALSE	0	NA	2018-02-14 19:30:41	TRUE	NA	9638580497850040...
4	Audio de la clase de miner<U+00ED>a de datos. Los ...	FALSE	0	NA	2018-02-14 19:28:14	FALSE	NA	9638574331343544...
5	V<U+00ED>deo de La miner<U+00ED>a de Texto (T...	FALSE	0	NA	2018-02-14 19:23:03	FALSE	NA	9638561275649638...
6	#UFCxFOX Disfrutando el Main Card con mi familia de...	FALSE	0	NA	2017-12-31 03:21:17	FALSE	NA	9473066374795468...
7	#UFCxFOX Un saludo desde M<U+00E9>rida Venezu...	FALSE	0	NA	2017-12-31 01:38:30	FALSE	NA	9472807709383516...
8	ANC regular<U+00E1> uso de redes sociales para sa...	FALSE	0	NA	2017-08-29 14:40:06	FALSE	NA	9025413748468899...
9	#MoneyFightxFOX Gana Macgregor. Para Los seguido...	FALSE	2	NA	2017-08-27 03:34:43	FALSE	NA	9016491476605255...
10	#MoneyFightxFOX un saludo desde M<U+00E9>rida ...	FALSE	0	NA	2017-08-27 03:16:18	FALSE	NA	9016445154249932...
11	Twitter Apps https://t.co/xWjy2VjQTa v<U+00ED>a ...	FALSE	0	NA	2017-08-07 08:49:23	FALSE	NA	8944805817524879...
12	@DolarToday C<U+00F3>mo es posible que lo dejar...	FALSE	0	DolarToday	2017-05-13 19:32:39	FALSE	863471674871697415	8634771054468546...

Figura 5.4: Data frame del timeLine de los estudiantes

- Filtrado de Tweets:** Los tweets capturados, son filtrados utilizando la librería tm de minería de texto de R. Para ese filtrado, fueron utilizadas como palabras clave (de filtrado) las actividades, estrategias, instrumentos educativos y herramientas de aprendizaje, definidas en el modelo de Felder Silverman de estilos de aprendizaje (ver apéndice E). Este procedimiento permite obtener las actividades académicas preferidas por los estudiantes en la red social (ver Figura 5.5, y apéndice C con los scripts usados en este caso).

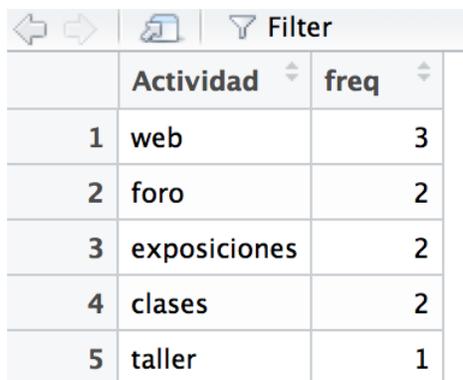
```

113
114
115 #Filtro las actividades de aprendizaje
116 |
117 dm=filter(dm, Actividad %in% c("web", "taller", "foro", "video","exposiciones","clases", "fotos"))
118
119 #creo un csv con las actividades de aprendizaje
120 write.csv(dm,file="actividadesCorpus.csv")
121
122

```

Figura 5.5: Filtro de las actividades de aprendizaje

En particular, esta fase culmina con la determinación de la frecuencia con la que cada estudiante realiza las diferentes actividades académicas establecidas en el modelo de Felder Silverman, en las redes sociales (ver Figura 5.6).



	Actividad	freq
1	web	3
2	foro	2
3	exposiciones	2
4	clases	2
5	taller	1

Figura 5.6: Frecuencia de actividades de un estudiante.

E1P5: Unir las dos listas de actividades anteriores para obtener Lista 1

En este paso se unen las listas de actividades obtenidas en los dos pasos anteriores, la lista obtenida desde Moodle, con la lista de actividades de los estudiantes obtenida en la plataforma de red social twitter, para obtener una nueva lista de actividades más general, denominada todosLista_1.txt.

E1P6: Obtención de lista de actividades ofrecidas por la materia Lista 2

Esa lista de actividades es obtenida con una consulta SQL a la base de datos de Moodle (ver Figura 5.7), para obtener la lista de actividades 2 (todosLista_2.txt).



```
select * from moodle19.mdl_assignment
where course=57063
```

Figura 5.7: Consulta de las actividades académicas que se deben realizar en el curso de computación

La tabla 5.3 muestra las tablas de la base de datos involucradas en la consulta, y la descripción de los campos utilizados de ellas.

Tablas Involucradas	Descripción
moodle19.mdl_assignment	Contiene las actividades de aprendizaje ofrecidas por las materias en la plataforma moodle.
Campos de la tabla	Descripción
mdl_assignment.course	Contiene el identificador de la materia dentro de la tabla.
mdl_assignment.name	Contiene el nombre de la asignación de actividad en la tabla, por ejemplo: Examen, Tarea, Ensayo etc.
mdl_assignment.description	Contiene la descripción de cada una de las actividades que serán utilizadas en el curso.

Tabla 5.3: Tablas y campos utilizados en las consultas SQL.

E1P7: Realizar la mezcla entre las actividades de las lista anteriores para obtener Lista 3

En este paso es realizada la mezcla de las listas de actividades obtenidas en E1P6 y E1P5, para obtener una nueva lista de actividades. Este proceso se genera con una línea de comandos sencilla, que incorpora el contenido de las dos listas en una lista final, denominada todosLista_3.

E1P8: Determinar la tabla de frecuencias de las actividades realizadas en cada dimensión del modelo de Felder Silverman, de la lista anterior

En este paso se obtiene la lista de frecuencia de las actividades realizadas de cada dimensión del modelo de Felder Silverman, de la lista anterior almacenada en el archivo todosLista_3, y el resultado es guardado en dos archivos, denominados

frecTodosLista_3.txt y frecTodosLista_3.csv. El código realizado para este proceso se puede apreciar en el apéndice E. El resultado se muestra en la siguiente tabla:

1	2	3	4	5	6	7	8
Sensitivo	Intuitivo	Visual	Verbal	Activo	Reflexivo	Secuencial	Global
6	2	16	49	9	4	0	13

Tabla 5.4: Frecuencias de las actividades realizadas en cada dimensión del modelo de Felder Silverman

EIP9: Determinar el estilo de aprendizaje del curso a partir de la lista de frecuencias de actividades general

Con la lista de actividades ya definida, se puede determinar el estilo de aprendizaje utilizado por los mejores estudiantes del actual curso, el cual es el estilo de aprendizaje inicial usado por SaCI para organizar sus actividades. Para ello, utilizamos el script del apéndice D y el resultado se puede apreciar en la figura 5.8.

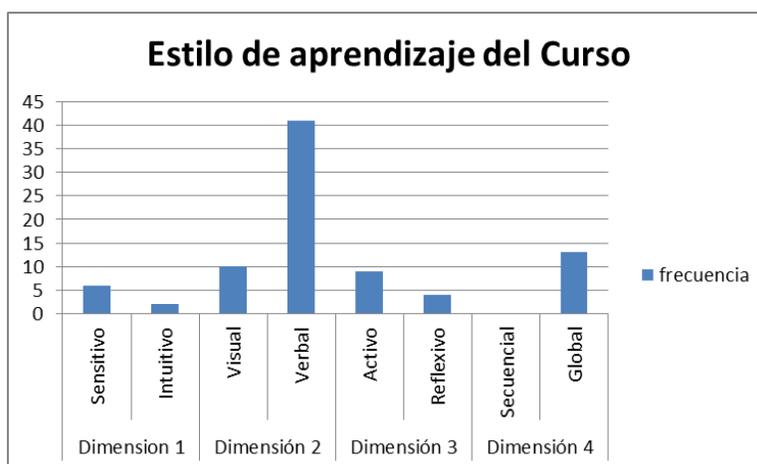


Figura 5.8: Estilo de aprendizaje del Curso de Computación 5284 UTPL

En este caso, se puede apreciar que el estilo de aprendizaje que predomina es:

Sensitivo, Verbal, Activo y Global

5.1.2 Escenario de prueba Nro. 2.

Objetivo

Determinar el estilo de aprendizaje adecuado de la carrera de Derecho. En este caso se conoce los estilos de aprendizaje de sus estudiantes, y se quiere con ello determinar el estilo de aprendizaje ideal para la carrera de Derecho en la UTPL.

Características del contexto

Se tratan los mejores estudiantes de la carrera de Derecho, de los cuales se tomaron los datos de la plataforma Moodle. En este caso, a diferencia del anterior escenario, ya se poseen los estilos de aprendizaje como punto de partida, y queremos determinar el estilo de aprendizaje ideal de la carrera, obtenido a partir de las actividades académicas que realizan los mejores estudiantes, sus estilos, y sus comportamientos en las redes sociales.

Aplicación del Macro-Algoritmo

Tarea	Paso	Detalle	Tarea de Minería de Datos	Herramientas usada	Producto
1	E2P1	Obtención de los datos de los estudiantes SaCI	-	SQL	Lista de estudiantes
1	E2P2	Agrupar estudiantes por rendimiento académico.	Agrupamiento de los estudiantes según su rendimiento histórico	SQL y SPSS (para las áreas de agrupamiento).	Estudiantes agrupados en: Alto Rendimiento
2	E2P3	Obtención de lista de actividades académicas de los estudiantes de mejor rendimiento, basada en el modelo Felder Silverman.		SQL	Lista de actividades académicas actuales de los estudiantes
2	E2P4	Obtención de la lista de actividades realizadas en las redes sociales, por los estudiantes de mejor	Minería de texto en Twitter. Filtro de actividades	R (Ver Apéndice C)	Actividades y Frecuencia

		rendimiento, basada en el modelo Felder Silverman	según estilos de aprendizaje de los estudiantes (ver apéndice E)		
2	E2P5	Unir las dos listas de actividades			Lista de actividades académicas y Sociales de los estudiantes
3	E2P6	Obtención de lista de actividades ofrecidas por las materias de las carreras de Derecho		SQL	Lista de las actividades que en ese curso de computación se deben realizar.
3	E2P7	Realizar la mezcla entre las actividades de las listas anteriores			Lista de todas las actividades identificadas en los pasos anteriores
4	E2P8	Determinar la tabla de frecuencias de las actividades de cada dimensión del modelo de modelo Felder Silverman	Calcular el número de actividades realizadas en cada dimensión del modelo Felder Silverman, contenidas en la tabla de frecuencias	R (ver Apéndice D)	Tabla General de Frecuencias de actividades:
	E2P9	Determinar el estilo de aprendizaje de la carrea de derecho a partir de la lista de frecuencias de actividades general	Uso de la tabla de frecuencias para obtener el estilo de aprendizaje	R (ver Apéndice D)	

Tabla 5.5: Macro-Algoritmo para el segundo escenario

E2P1: Obtención de datos de los estudiantes SaCI

En este caso, se recuperan desde la base de datos todos los estudiantes de la carrera de derecho, con una consulta SQL (ver Figura 5.1), pero ahora la categoría de las materias es 5223, que corresponde a todos los cursos de la carrera Derecho.

E2P2: Agrupar estudiantes por rendimiento

Con el archivo de la consulta obtenida en el paso anterior, se procede a agrupar a los estudiantes con mejor rendimiento en la carrera. El resultado de la agrupación se puede apreciar en la Figura 5.9, correspondiente al algoritmo de K-means, que muestra en este caso que los alumnos de alto rendimiento corresponden al cluster 3. Son aquellos cuyo promedio de notas es superior a 91.

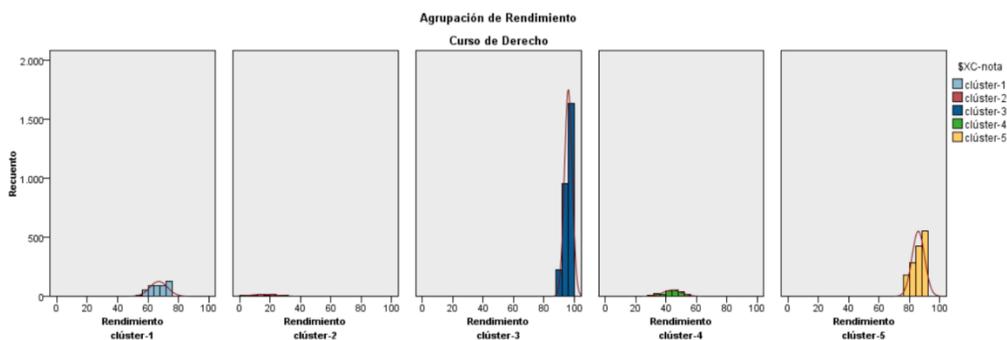


Figura 5.9: Agrupamiento K-means del curso de Derecho.

E2P3: Obtención de lista de actividades académicas desarrolladas en la carrera de Derecho por los estudiantes de mejor rendimiento.

A continuación es realizada la consulta en la plataforma moodle de SaCI, de la lista de actividades académicas realizada por los estudiantes de mejor rendimiento (dichas actividades, son las establecidas por el modelo de Felder Silverman). El resultado se guarda es una archivo llamado actividadesMoodle5223.txt

E2P4: Obtención de la lista de actividades realizadas por los estudiantes de mejor rendimiento en las redes sociales

En este paso del algoritmo, se utilizó nuevamente el API de Twitter para analizar el comportamiento de los mejores estudiantes de la carrera de Derecho en esa red social. Al final, se obtienen un archivo con todas las actividades según el modelo de Felder Silverman, realizadas por los mejores estudiantes de Derecho en Twitter, denominado todosRD_5223.txt.

E2P5: Unir las dos listas de actividades (P3, P4) para obtener Lista 1

En este paso se unen las dos listas de actividades anteriores, para obtener una lista de actividades más general, denominada todosLista_1.txt

E2P6: Obtención de lista de actividades académicas que se deben realizar en las materiales de la carrera de Derecho, llamada Lista 2

Para ello, se hace una consulta SQL al Moodle de SaCI, que recupera la lista de actividades académicas (establecidas en el modelo de Felder Silverman) que se deben realizar en las materias de la carrera de Derecho. En este paso se obtiene dicha lista de actividades (todosLista_2_5223.txt).

E2P7: Realizar la mezcla entre las actividades de las listas de los dos últimos pasos, para obtener Lista 3

En este paso es realizada la mezcla de las listas de actividades académicas obtenidas en los dos últimos pasos, denominada todosLista_3_5223.txt.

E2P8: Determinar la tabla de frecuencia de las actividades académicas contenidas en la lista anterior

En este paso, se obtiene la frecuencia de las actividades académicas realizadas, definidas según el modelo de Felder Silverman, y el resultado es guardado en dos archivos, denominados frecTodosLista_3_5223.txt y frecTodosLista_3_5223.csv, de donde será definido el estilo de aprendizaje actual de la carrera de Derecho (ver Tabla 5.6).

1	2	3	4	5	6	7	8
Sensitivo	Intuitivo	Visual	Verbal	Activo	Reflexivo	Secuencial	Global
82	60	101	63	95	63	84	61

Tabla 5.6: Tabla de Frecuencias de las actividades general

E2P9: Determinar el estilo de aprendizaje de la carrera de Derecho

Con la lista de actividades obtenida, se procedió a determinar el estilo de aprendizaje actual de los estudiantes de la carrera de Derecho (el script para calcularlo está en el apéndice D). El resultado se puede apreciar en la figura 5.10.

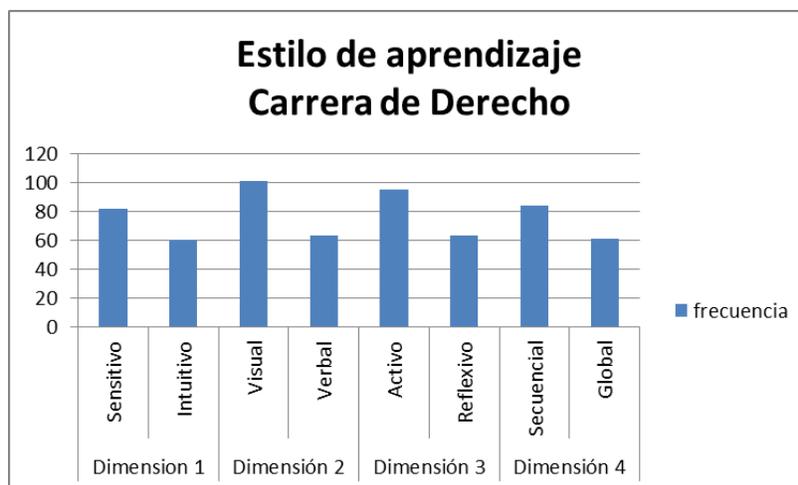


Figura 5.10: Estilo de aprendizaje de la categoría 5223 de Derecho UTPL

En la Figura 5.10, se puede apreciar que el estilo de aprendizaje ideal en la carrera de Derecho es:

Sensitivo, **Visual**, Activo y **Secuencial**

5.2 Análisis de Resultados

La incorporación de Analítica Social de Aprendizaje en los ciclos autonómicos de tareas en SaCI, permite construir un mejor perfil de un estudiante de un curso o de una carrera. En general, permite incorporar datos externos para construir modelos y patrones de datos más completos que ayudan a lograr los objetivos. En muchos casos en los que se han construido estilos de aprendizaje, en sistemas académicos por sí solos o por la aplicación del test del modelo Felder Silverman, no han sido suficientes [23]. Con este trabajo creamos el

precedente de que con la aplicación de ciclos autonómicos basados en tareas de analítica social de aprendizaje, es posible generar estilos de aprendizaje más completos, basados en el comportamiento de los actores dentro de un salón de clases inteligente y fuera de él, en particular, en las redes sociales.

Para los dos escenarios desarrollados en el capítulo 5, el objetivo fue descubrir el estilo de aprendizaje adecuado en un curso o en una carrera dada, aplicando el ciclo autonómico de tareas descrito en el capítulo 4. Este ciclo funciona a partir de la construcción de los perfiles que resultan de las actividades que realizan los estudiantes en la plataforma Moodle, las actividades que realizan estos mismos estudiantes en las redes sociales, y las actividades que son requeridas por el (los) curso(s). La diferencia en cada escenario radicó en dos cosas. Por un lado, que para el primero el estilo de aprendizaje de los estudiantes del curso era desconocido, mientras que en el segundo no lo era. La segunda cosa es que para el primer caso se determinó el estilo de aprendizaje para un curso dado, y para el segundo de una carrera completa. En este sentido, se instanció el ciclo autonómico de dos maneras diferentes (cada escenario), para la obtención de un estilo de aprendizaje adecuado.

Ahora bien, en ambos escenarios se construye un estilo de aprendizaje por consenso, el cual es el socialmente adecuado para un curso o una carrera dada. Ese estilo de aprendizaje es el que debe usar SaCI para organizar sus actividades, buscar y presentar los objetos de aprendizaje, entre otras cosas.

5.2.1 Resultados Escenario 1

En este escenario se determinó que el estilo de aprendizaje adecuado para el curso de Computación es el **“Sensitivo, Verbal, Activo y Global”**. Para obtener el resultado, fue necesario extraer y agrupar en categorías a los estudiantes de ese curso, este paso se logró con la aplicación de la tarea de Analítica de Datos de agrupación, basado en la distancia. En particular, se usó el algoritmo K-Means, cuyo resultado fue de tres grupos (Alto rendimiento, Medio Rendimiento y Bajo Rendimiento). Luego, se obtuvieron las actividades realizadas por el grupo de alto rendimiento en la plataforma Moodle, y en las Redes Sociales (en este caso, twitter). Para ello, se realizó minería de texto en ambos contextos, para buscar por palabras claves información sobre esas actividades. Después, se

generó una tabla de frecuencias de aparición de esas actividades, para construir el perfil del estilo de aprendizaje adecuado para el curso.

El estilo de aprendizaje obtenido muestra que los contenidos que obtuvieron mayor éxito son los que corresponden a: laboratorios con experimentos y talleres (Sensitivo), Charlas, foros y discusiones en grupo (Verbal), ejercicios y ensayos (Activos), resolución de problemas y métodos de casos (Global). Esta información le sirve a los agentes de SaCI involucrados, como el entorno virtual de aprendizaje VLE y el sistema recomendador SR, para incorporar contenidos y actividades adecuadas en el proceso de aprendizaje del curso.

Algunas de las tareas de este ciclo autónomico pueden ser reutilizadas en otros ciclos autónomicos para alcanzar otros objetivos. Por ejemplo, la tarea de agrupación que agrupa los estudiantes por rendimiento, puede ser reusada por el ciclo autónomico que trata de evitar la deserción estudiantil. En ese caso, el grupo utilizado sería el de los estudiantes de bajo rendimiento, o los estudiantes con necesidades puntuales.

5.2.2 Resultados Escenario 2

En este escenario se determinó que el estilo de aprendizaje adecuado para la carrera de Derecho es el **“Sensitivo, Visual, Activo y Secuencial”**. Para obtener ese resultado, fue necesario identificar a los estudiantes de la carrera de derecho. En ese escenario, la aplicación de la tarea de Analítica de Datos de agrupación permitió identificar los mejores estudiantes de la carrera. Luego, se obtuvieron las actividades realizadas por dichos estudiantes en la plataforma Moodle y en las Redes Sociales (twitter). Después, se determinó la tabla de frecuencias con las actividades que fueron realizadas, para construir el perfil de estilo de aprendizaje de la carrera.

El estilo de aprendizaje obtenido muestra que los contenidos que obtuvieron mayor éxito son los que corresponden a: laboratorios con experimentos y talleres (Sensitivo), el uso de Videos, Exposiciones, figuras y gráficas (Visual), ejercicios y ensayos (Activos), y diagramas de flujo (Secuencial). Igual que en el anterior escenario, esta referencia sirve para que los agentes de SaCI, como el agente del entorno virtual de aprendizaje VLE y el

agente recomendador SR, determinen que incorporar como contenido y actividades en el proceso de enseñanza de la carrera.

Particularmente, en este escenario se puede observar que el estilo de aprendizaje inicial de la carrera “**Sensitivo, Verbal, Activo Global**”, es diferente al estilo de aprendizaje obtenido “**Sensitivo, Visual, Activo y Secuencial**”. Así, vemos que el ciclo autonómico puede determinar cómo va cambiando el estilo de aprendizaje en un curso/carrera, lo cual es de gran importancia para la gestión del proceso de aprendizaje por parte de SaCI, de tal manera que el mismo tenga los contenidos y actividades acorde a las reales necesidades de los estudiantes.

5.2.2 Comparación con trabajos similares

En otros trabajos, como el proyecto AARTIC [5], el sistema ayuda a los estudiantes a entender los conceptos usados. Ahora bien, en el proyecto AARTIC lo hace siguiendo el esquema de aprendizaje con el cual fue programado el agente, sin tomar en cuenta las necesidades cambiantes que pueda tener un estudiante para lograr el aprendizaje, ni el perfil de aprendizaje o limitaciones que pueda encontrar en el proceso. Existen situaciones en las que el estilo de aprendizaje no depende de las preferencias del estudiante, sino de variables externas, como la velocidad del internet o fallas de corriente eléctrica. En este tipo de situaciones, el estudiante puede tender a ser más verbal que visual para aprovechar el ancho de banda (es más fácil descargar un texto que ver un video). Este tipo de situaciones deben ser contempladas al momento de definir el estilo de aprendizaje, lo cual es posible desde nuestros ciclos autonómicos, al poder detectar esas situaciones a través de las redes sociales.

En proyectos como ISABEL [3], los estudiantes son divididos en grupos con perfiles similares, y cada grupo es gestionado por un agente tutor. Además, hay un agente profesor asociado a cada grupo. Esta es una interesante propuesta, sin embargo, en esa propuesta no se aplica analítica de aprendizaje para aprovechar los patrones obtenidos de los datos, y así mejorar el proceso de enseñanza. Tampoco se actualizan los perfiles de los estudiantes en los grupos.

La tabla 5.7 presenta una comparación general entre varios trabajos vinculados a proponer ambientes inteligentes de aprendizaje (como SaCI), considerando si esos trabajos usan tareas de AdD de aprendizaje (LA) y tareas de AdDS de aprendizaje (SLA), si usan ciclos autónomos de tareas de AdD (CA), si son para procesos online, y si utilizan los estilos de aprendizaje de los estudiantes (EA).

Vemos que solo nuestra propuesta y [22], usan SLA para generar conocimiento con los datos del entorno del ambiente inteligente; y solo tres de los trabajos, incluyendo el nuestro, usan LA para generar conocimiento con los datos del AmI. Además, solo nuestra propuesta plantea un CA para automatizar el proceso de generación de conocimiento desde los datos, Finalmente, todos los trabajos son para AmIs, y solo nuestra propuesta y [3] utilizan los estilos de aprendizaje de los estudiantes.

Propuestas	LA	SLA	CA	OL	EA
SaCI con CA de tareas de AdD	✓	✓	✓	✓	✓
Proyecto AARTIC [5]	✓	○	○	✓	○
Proyecto ISABEL [3]	○	○	○	✓	✓
e-Learning [2]	✓	○	○	✓	○
Proyecto Chamilo [22]	✓	✓	○	✓	○

Tabla 5.7: Cuadro comparativo entre SaCI y otras propuestas (Leyenda: LA=Learning Analytics, SLA=Social Learning Analytics, CA=Ciclos Autónomos, OL=Online EA=Estilos de Aprendizaje)

Capítulo 6

Conclusiones y Recomendaciones

6.1 Conclusiones

En este trabajo se incorporó la Analítica Social de Aprendizaje, al proceso de generación de conocimiento para los salones de clase inteligentes. La propuesta fue desarrollada, incorporando el concepto de ciclo autonómico de tareas de analítica de datos, el cual describe la arquitectura de computación autonómica para el alcance de los objetivos claves de los salones de clase inteligentes. Además, se caracterizaron diferentes objetivos vitales de los salones de clase inteligentes, que pueden ser alcanzados con dichos ciclos autonómicos. Finalmente, se especificaron las tareas de analítica de datos para cada ciclo.

Los ciclos autonómicos especificados para SaCI, fueron para determinar el paradigma de aprendizaje adecuado para un curso, para determinar el recurso educativo ideal para un estudiante, para identificar los estudiantes con necesidades específicas para poder ayudarlo, y por último, para evitar la deserción estudiantil. Futuros estudios podrán definir nuevos ciclos autonómicos.

Las tareas de analítica de datos social y clásica, fueron desarrolladas por módulos independientes y reutilizables, lo que permite que otros ciclos autonómicos sean capaces de incorporarlas en sus estructuras. En este sentido, algunas tareas fueron desarrolladas mediante Scripts reutilizables en R y en clases de Java de fácil acceso. El prototipo de esas tareas se encuentra disponible en los apéndices de este trabajo.

Para realizar la experimentación en SaCI, fue seleccionado el CA **“Determinar el paradigma de aprendizaje adecuado para un curso”**. Fueron instanciadas cada una de las tareas del ciclo, y el CA fue probado en dos escenarios. Un escenario tenía como objetivo obtener el estilo de aprendizaje de un curso (en nuestro caso, un curso de computación), y otro escenario actualizar el perfil de aprendizaje de una carrera (en nuestro caso, de la carrera de Derecho).

De acuerdo a la revisión de la literatura, no hay precedentes de obtención del estilo de aprendizaje de un curso, de una carrera, o de un estudiante, de forma automática, basado en el comportamiento de dichos estudiantes en diferentes contextos (aulas de clases, redes sociales, etc.). Así, el CA de SaCI permite la obtención de estilos de aprendizaje en diferentes contextos (de un curso cuando SaCI aún no lo conoce, actualizar el de una carrera, etc.), para lo cual usa diferentes tareas de analítica de datos. Algunas conclusiones generales de este caso de estudio son:

- El ciclo Autonomico de tareas puede obtener el estilo de aprendizaje para un grupo de cursos (por ejemplo, una carrera), para un curso, o para un estudiante. El resultado depende del corpus de datos suministrado como entrada.
- Se pueden utilizar plataformas externas para complementar el objetivo. En este caso se usó Twitter como plataforma social externa, ya que esta red social estaba siendo utilizada por algunos profesores y alumnos de los salones de clase inteligentes para interactuar. Ahora bien, podría usarse cualquier otra red como Facebook, Instagram entre otras.
- El ciclo autonomico de tareas representa una mejora considerable en el proceso de enseñanza y aprendizaje en un SaCI, debido a que es capaz de generar conocimiento constantemente, desde fuentes internas y externas a SaCI, útil para mejorar dicho proceso.
- El ciclo es capaz de incorporar tareas de AdDS, lo cual le permite establecer mejores perfiles de estilos de aprendizaje.

6.2 Recomendaciones

Es importante que todos los estudiantes y profesores registrados en SaCI, cuenten con sus respectivas cuentas de redes sociales registradas en la plataforma. Este objetivo puede alcanzarse mediante la creación de un nuevo ciclo autonómico que este descubriendo las cuentas en las redes sociales de los usuarios de SaCI.

Por otro lado, el modelo de estilos de aprendizaje utilizado en esta tesis fue el de Felder Silverman, sin embargo, existen otros modelos para establecer los paradigmas de aprendizaje. En este sentido, sería interesante hacer los ajustes para que puedan ser incorporados nuevos modelos de estilos de aprendizaje al macro algoritmo. Esto permitiría la obtención de un perfil más completo, incorporando nuevas actividades que no son contempladas en este modelo, tales como las relaciones interpersonales del estudiante previstas en el Modelo de Grasha y Riechman [21]. Para ello, ese modelo incluye una dimensión sobre la actitud del alumno hacia el aprendizaje (participativo/elusivo), otra sobre su actitud hacia los compañeros y maestros (competitivo/colaborativo), y otra sobre sus reacciones hacia los procesos didácticos (dependiente/independiente).

También, los estilos de aprendizaje que van teniendo los estudiantes y los cursos, podrían ser guardados en un histórico de datos. Esto permitirá buscar patrones de comportamiento a partir de esos datos. Otro aspecto importante a estudiar es la integración de todos los CA en un contexto real, para determinar la sinergia que los mismos producirían actuando simultáneamente.

Finalmente, es necesario generar métricas pedagógicas para evaluar el impacto de estas herramientas tecnológicas en los entornos de aprendizaje. Además, es fundamental generar experimentos en entornos reales, que permitan medir en dichos contextos el aporte de estas tecnologías en los procesos de enseñanza-aprendizaje.

Referencias Bibliográficas

- [1] Cook, D.J., Augusto, J.C., Jakkula, V.R., “Ambient intelligence: Technologies, applications, and opportunities”, *Pervasive and Mobile Computing*, Vol. 5, pp. 277-298, (2009).
- [2] Aguilar, J., Valdiviezo, P., Cordero, J., Sanchez, M., “Conceptual Design of a Smart Classroom Based on Multiagent Systems”, *Intl Conf. Artificial Intelligence ICAI'15*, (2015).
- [3] Garruzzo, S., Rosaci, D., “Isabel: A multi agent e-learning system that supports multiple devices”, *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, (2007).
- [4] Y. Shi, W. Qin, Y. Suo, X. Xiao., “Smart classroom: Bringing pervasive computing into distance learning”, *Handbook of Ambient Intelligence and Smart Environments*, Springer, pp. 881-910, (2010).
- [5] Mhiri, F., Ratté, S., “AARTIC: development of an intelligent environment for human learning”, *Proceedings of the 14th annual ACM SIGCSE conference on Innovation and technology in computer science education (ITiCSE 2009)*, (2009).
- [6] Calderón, H., Guerrero, E., Godoy, C., “Ambientes Inteligentes en Contextos Educativos: Modelo y Arquitectura”, *Research in Computing Science*. Vol. 77, (2014).
- [7] Sánchez, M., Aguilar, J, Cordero, J., Valdiviezo P., “A Smart Learning Environment based on Cloud Learning”, *International Journal of Advanced Information Science and Technology*, Vol. 39, No. 39, pp. 39-52, (2015).
- [8] Aguilar, J., Cordero, J., Buendía, O., “Specification of the Autonomic Cycles of Learning Analytic Task for a Smart Classroom”, *Journal Of Educational Computing Research*, (2017).

- [9] Rebeca, F., “Social Learning Analytics”, Institute of Educational Technology the open university UK, Vol. 4, pp. 161-176, (2015).
- [10] Sánchez, M., Aguilar, J., Cordero, J., Valdiviezo, P., Barba, L., Chamba, L., “Cloud computing in smart educational environments: application in learning analytics as service”, New Advances in Information Systems and Technologies. Advances in Intelligent Systems and Computing, Vol. 444. Springer, (2016).
- [11] Aguilar, J., Cordero, J., Valdiviezo, P., “Basic features of a reflective middleware for intelligent learning environment in the cloud”, Asia-Pacific Conference on Computer Aided System Engineering, pp. 1–6, (2015b).
- [12] Ferguson, R., Shum, S.B., “Social learning analytics: five approaches”, Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 23-33, (2012).
- [13] Aguilar, J., Buendía, O., Moreno, K., Mosquera, D., “Autonomous Cycle of Data Analysis Tasks for Learning Processes”, Valencia-García R., Lagos-Ortiz K., Alcaraz-Mármol G., del Cioppo J., Vera-Lucio N. (eds) Technologies and Innovation. CITI 2016. Communications in Computer and Information Science, Vol. 658. Springer, (2016).
- [14] Gonzalez, A., Buendía, O., Aguilar, J., Cordero, J., Rodriguez, T., “Competencies as Services in the Autonomic Cycles of Learning Analytics Task for a Smart Classroom”, Technologies and Innovation Third International Conference, CITI 2017, (Proceedings), pp. 211-226, Guayaquil, Ecuador, (2017).
- [15] Wikipedia., “Learning Data”, recuperado el 12 de Noviembre de 2017, de https://es.wikipedia.org/wiki/Learning_analytics#cite_note-29,20189, (2017).
- [16] Aguilar, J., Dos Santos, R., Gonzalez, A., Belen, M., Moreno, K., Rodriguez, T., Tenssaca, G., ”Introducción a la Minería Semántica”, Primera Edición, Universidad Nacional Experimental del Táchira, (2018).
- [17] Lin, A., “Conceptual Model for Business-Oriented Management of Web Services”, Proceedings of the 6th WSEAS Int. Conf. on Software Engineering, Parallel and Distributed Systems, pp. 80-85, (2007).

[18] Rangel, Carlos R., “Ontologías desde el punto de vista de Minería Ontológica y del Paradigma de Arquitecturas Orientadas a Servicios”, Universidad de Los Andes, (2015).

[19] Rangel. R., “Integración de Ontologías desde el punto de vista de Minería Ontológica y del paradigma de arquitecturas orientadas a servicios”, Tesis de grado de Maestro en ciencias de la computación. Instituto Politécnico Nacional Centro de Investigación en Computación, (2015).

[20] Rangel, C. Pacheco, F., Aguilar, J., Cerrada, M., Altamiranda, J., “Metodología para identificar donde extraer conocimiento en una organización extendido”, Universidad de Los Andes, (2015).

[21] Scielo Cientific Electronic Library online., Student Learning Styles, por Grasha A. F. y Riechmann, S. W., recuperado el 10 de Enero de 2018, de <http://www.scielo.org.co/pdf/rcde/n64/n64a04.pdf>, (2012).

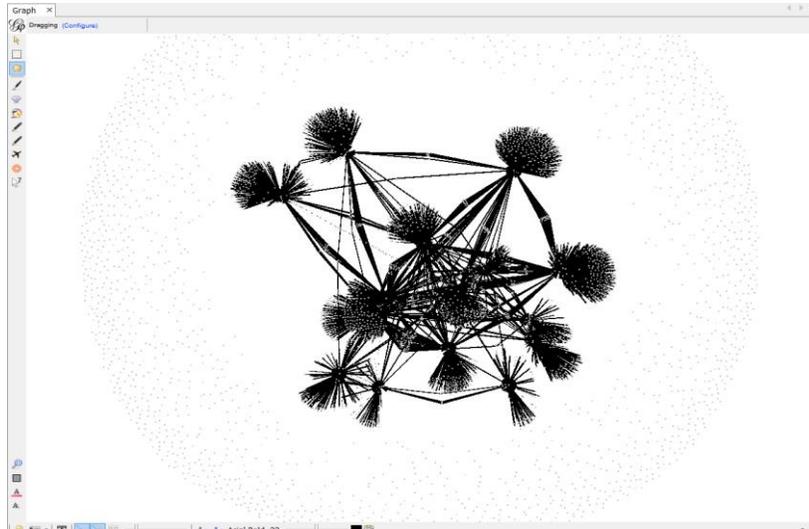
[22] SlideShare., Introduccion al Proyecto Chamilo, recuperado el 10 de Enero de 2018, de <https://es.slideshare.net/chamiluda/introduccion-al-proyecto-chamilo-lms-qu-es-chamilo>, (2017).

[23] Aguilar, J., Fuentes, J. Moreno, K. Dos Santos, R., Portilla, O., Altamiranda, J. Hernández, D., “Computational Platform for the Educational Model based on the Cloud Paradigm”, Proceeding of the 45th Annual Frontiers in Education (FIE) Conference, 2399-2407, (2015).

Apéndices

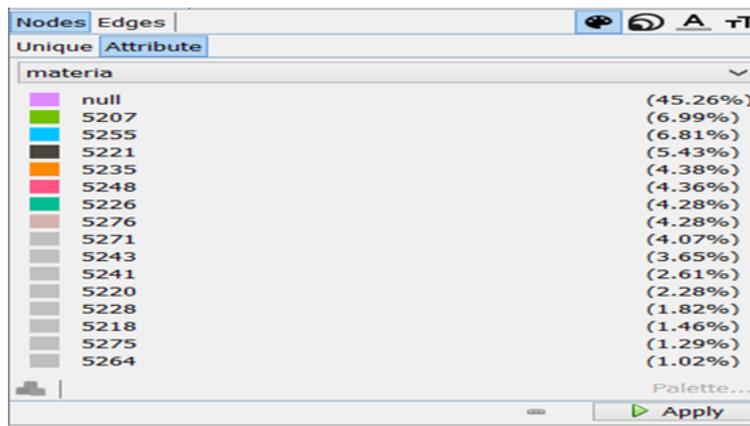
Apéndice A. Mediciones para obtener los estudiantes de alto rendimiento

El grafo a continuación, representa todos los alumnos inscritos, agrupados por categorías (carreras) a las que pertenece cada curso.



Grafo general de estudiantes inscritos SaCI algoritmo Yifan Hu

Categorías con mayor número de alumnos inscritos y por lo tanto, las de más relevancia



Distribución por categorías grafo de estudiantes inscritos

Apéndice B. Programa En Java para Clasificar los estudiantes por rendimiento

Clase Estudiante

```
package csv.clasificador;
public class Estudiante {

    //Variables
    private String userId;
    private String materia;
    private String finalGrade;
    private String rendimiento;

    //Constructor de la clase
    public Estudiante(String userId, String materia, String finalGrade, String rendimiento) {
        this.userId = userId;
        this.materia = materia;
        this.finalGrade = finalGrade;
        this.rendimiento = rendimiento;}
    //métodos get y set
    public void setUserId(String userId) {
        this.userId = userId;}

    public String getMateria() {
        return materia;}

    public void setMateria(String materia) {
        this.materia = materia;}
    public String getFinalGrade() {
        return finalGrade;}

    public void setFinalGrade(String finalGrade) {
        this.finalGrade = finalGrade;}

    public String getRendimiento() {
        return rendimiento;}
```

Clase Principal

```
package csv.clasificador;
import java.io.FileNotFoundException;
import java.io.IOException;
import java.util.ArrayList;
import java.util.List;
import com.csvreader.CsvReader;
import com.csvreader.CsvWriter;
import java.io.File;
import java.io.FileWriter;

public class ClasificarEst {
    private String clase,rendimiento;

    public static void main(String[] args) {
        File archivo = new File("c://tesis2018//estCursoClase.csv");
        if (archivo.exists()) {archivo.delete();}
        try {
            List<Estudiante> usuarios = new ArrayList<Estudiante>();
            CsvReader estudiantes_importados = new CsvReader("c://tesis2018//estCurso.csv");
            estudiantes_importados.readHeaders();
            //genera el archivo en la carpeta del proyecto
            CsvWriter writer = new CsvWriter(new FileWriter("estCursoClase.csv", true), ',');
            //bucle que asigna los valores del archivo leído a las variables
            while (estudiantes_importados.readRecord()) {
                String userId = estudiantes_importados.get(0); String materia =
estudiantes_importados.get(1); String finalGrade = estudiantes_importados.get(2); String rendimiento = "";
                //clasificación de los rendimientos
                if (Double.parseDouble(finalGrade) < 80) { rendimiento = "No";
                } else { rendimiento = "Si";
                }
                usuarios.add(new Estudiante(userId, materia, finalGrade, rendimiento));
            } estudiantes_importados.close();

            //bucle para escribir los resultados en el archivo de salida
            for (Estudiante us : usuarios) {
                System.out.println(us.getUserId() + " " + us.getMateria() + " "
                + us.getFinalGrade() + " " + us.getRendimiento());

                writer.write(us.getUserId());writer.write(us.getMateria());
                writer.write(us.getFinalGrade()); writer.write(us.getRendimiento());
                writer.endRecord(); writer.flush();
            }
        } catch (FileNotFoundException e) { e.printStackTrace();
        } catch (IOException e) {e.printStackTrace();
        }
    }
}
```



```

#Universidad de Los Andes
#Postgrado en Computación
#Script para obtener el timeline de los estudiantes de la categoría de computación UTPL
#Realizado por Omar Buendía
#Página 1

mach_text
# Se define la función clean.text Filtra palabras (evita URL, @usuarios, etc...)

clean.text = function(x)
{
  # tolower
  x = tolower(x)
  # remove rt
  #x = gsub("rt", "", x)
  # remove at
  #x = gsub("@\\w+", "", x)
  # remove punctuation
  x = gsub("[[:punct:]]", "", x)
  # remove numbers
  x = gsub("[[:digit:]]", "", x)
  # remove links http
  x = gsub("http\\w+", "", x)
  # remove tabs
  x = gsub("[ \\t]{2,}", "", x)
  # remove blank spaces at the beginning
  x = gsub("^ ", "", x)
  # remove blank spaces at the end
  x = gsub(" $", "", x)
  return(x)
}
mach_text = clean.text(mach_text)

# Transformamos la lista de textos en vectores y luego en una matriz
mach_corpus = Corpus(VectorSource(mach_text))

# Creamos una matriz, limpiamos los elementos que nos sobran (los artículos, y otras palabras sin sentido)
# Transformamos un Corpus en una matriz

tdm = TermDocumentMatrix(mach_corpus,
  control = list(removePunctuation = TRUE,
                 stopwords = c(stopwords("spanish"), stopwords("catalan"), "@"),
                 removeNumbers = TRUE, tolower = TRUE))
palabrasstop),
# definimos tdm as matrix
m = as.matrix(tdm)

# Contamos la frecuencia de las palabras de la lista de actividades de aprendizajes
# conteo de palabras en orden descendente
word_freqs = sort(rowSums(m), decreasing=TRUE)

# crea el data frame con las palabras y sus frecuencias
dm = data.frame(Actividad=names(word_freqs), freq=word_freqs)

# dibujamos una nube de plot wordcloud
wordcloud(dm$Actividad, dm$freq, random.order=FALSE, max.words=25, colors=brewer.pal(8, "Dark2"))

#Filtramos las actividades de aprendizaje

dm=filter(dm,
  Actividad %in% c("web", "taller", "foro", "video","exposiciones","clases", "fotos",
  "discusion en grupo","problemas", "juegos de roles","simulacion",
  "proyectos","talleres","laboratorio","demostracion","busqueda",
  "charla","metodos de casos","lecturas","ejercicios","cuestionario",
  "diagramas","figura","mapas conceptuales","imagen","grafico","texto",
  "diapositiva","tabla","autoevaluacion","examen","tutorial","portafolio",
  "experimento","conferencia","audio","pdf","articulo","utpl","conceptuales"))

mach_text

#creamos un csv con las actividades de aprendizaje
write.csv(dm,file="estudianteUTPL.csv")

#creamos un txt con los tweets del estudiante
write.table(mach_text, "estudianteUTPL.txt", sep=",")

#fin

```

Apéndice D. Script en R para determinar el estilo de aprendizaje del corpus total

```
#Universidad de Los Andes
#Postgrado en Computación
#Script para obtener Estilo de aprendizaje de
#de un corpus de datos en csv
#Realizado por Omar Buendia

library(twitteR)
library(tm)
library(wordcloud)
library(RColorBrewer)
library(dplyr)
library(RCurl)
library(data.table)

setwd("~/R/cuentasYeya")

dmf<-read.csv(file = "todos.csv",header = T,sep = ",")
#dimension 1
sensitivo=filter(dmf, Actividad %in% c("laboratorio","laboratorios", "experimento" ,"experimentos", "taller"))

intuitivo=filter(dmf, Actividad %in% c("busqueda","busquedas", "proyecto" ,"proyectos",
"tutorial","tutoriales"))

#dimension 2
Visual=filter(dmf, Actividad %in% c("exposicion","exposiciones", "demostracion" ,"demostraciones", "diagrama",
"diagramas","conceptual","mapa","mapas","grafica","grafico","graficos",
"video","videos","figura","figuras","imagen","imagenes","diapositivas",
"diapositiva","tabla","tablas","organigrama","organigramas"))

Verbal=filter(dmf, Actividad %in% c( "discusion","discusiones","grupo","grupos", "charla" ,"charlas","foro",
"foros","lectura","lecturas","examen","examenes","evaluacion","texto",
"textos","narrativa","narrativas","audio","audios"))

#dimension 3
activo=filter(dmf, Actividad %in% c("taller","talleres","ensayo","ensayos","laboratorio","laboratorios",
"ejercicio","ejercicios"))

reflexivo=filter(dmf, Actividad %in% c("ensayo","ensayos","simulacion","simulaciones","busqueda","busquedas",
"autoevaluacio","autoevaluaciones"))

#dimension 4
secuencial=filter(dmf, Actividad %in% c( "ejercicio", "ejercicios","diagramas-flujo","flujo","flujos"))
global=filter(dmf, Actividad %in% c("problemas","problemas","resolucion de problemas","metodos de casos" ))

sensitivo = colSums (select (sensitivo, contains ("freq")))
intuitivo = colSums (select (intuitivo, contains ("freq")))
visual = colSums (select (Visual, contains ("freq")))
verbal = colSums (select (Verbal, contains ("freq")))
activo = colSums (select (activo, contains ("freq")))
reflexivo = colSums (select (reflexivo, contains ("freq")))
secuencial = colSums (select (secuencial, contains ("freq")))
global = colSums (select (global, contains ("freq")))

boxplot(sensitivo,intuitivo,visual,verbal,activo,reflexivo,secuencial,global)

#fin
```

Apéndice E. Tabla de filtro de actividades con estilos de aprendizaje

TABLA DE FILTRO DE ACTIVIDADES CON ESTILOS DE APRENDIZAJE

Actividades / Estrategias / Instrumentos Educativos / Herramientas de Aprendizaje	Dimensión 1		Dimensión 2		Dimensión 3		Dimensión 4	
	Sensitivo	Intuitivo	Visual	Verbal	Activos	Reflexivos	Secuenciales	Globales
1 Discusión en Grupo				X				
2 Exposición			X					
3 Laboratorio	X							
4 Resolución de problemas								X
5 Talleres	X				X			
6 Charlas				X				
7 Ensayo	X					X		
8 Foro				X				
9 Simulaciones						X		
10 Demostraciones			X					
11 Búsqueda		X				X		
12 Metodos de Casos								X
13 Lecturas				X				
14 Ejercicios					X		X	
15 Examen Escrito				X				
16 Cuestionarios				X				
17 Evaluacion Oral				X				
18 Proyectos		X						
19 Experimento	X							
20 Tutoriales		X						
21 Autoevaluación						X		
22 Diagramas			X					
23 Mapa Conceptual			X					
24 Graficas			X					
25 Textos Narrativos				X				
26 Videos			X					
27 Audios				X				
28 Figuras			X					
29 Imágenes			X					
30 Diapositivas			X					
31 Tablas			X					
32 Diagrama de flujo							X	
33 Organigramas			X					

FILTRO DE ACTIVIDADES POR ESTILO DE APRENDIZAJE							
Sensitivo	Intuitivo	Visual	Verbal	Activos	Reflexivos	Secuencial	Global
Laboratorio	Búsqueda	Exposición	Discusión en Grupo	Taller	Ensayo	Ejercicios	Res. de Problemas
Experimento taller	Proyecto tutoriales	Demostraciones	Charlas	Ensayo	Simulaciones	Diag. de Flujo	Metodos de Casos
		Diagrama	Foros	ejercicios	Busquedas		
		Mapas conceptuale	Lecturas	laboratorio	Autoevaluacion		
		Gráficas	Examen escrito				
		Video	Evaluacion Oral				
		Figura	Texto narrativo				
		Imagen	Audios				
		diapositiva					
		Tabla					
		Organigrama					

Apéndice F. Modelo de Felder y Silverman

Felder Y Silverman aportan una clasificación para determinar los estilos de aprendizaje de los estudiantes, agrupándolos en 4 dimensiones. Estas dimensiones se forman a partir de preguntas clave para definir ¿Cómo aprendemos?

Dimensión 1: Información Percibida.

¿Qué tipo de información preferimos recibir cuando aprendemos?

Sensitivo: esta información va a ser captada de manera externa por los órganos de los sentidos. Implica observar y recopilar datos a través de los sentidos Por ejemplo: con un laboratorio, un experimento o un taller.

Intuitivo: la información va a ser captadas de manera interna por el sujeto a través de las percepciones que este tenga de la realidad de forma indirecta, por ejemplo: con la imaginación, la especulación, corazonadas, intuición etc. Prefieren realizar una búsqueda, hacer un proyecto de investigación.

Dimensión 2: Modalidad Sensorial.

¿Cómo se recibe la información?

Visual: si existe una tendencia o preferencia del estudiante a formatos gráficos, diagramas, mapas videos, figuras, diapositivas, etc.

Verbal: si existe una tendencia marcada del estudiante a actividades verbales como discusiones, charlas, foros, lecturas, textos y audios.

Dimensión 3 Preferencia para procesar la información

¿Qué Preferencia tiene el estudiante para procesar la Información?

Siendo Activo: haciendo algo con lo aprendido, tareas, practicas ejercicios.

Siendo Reflexivo: pensando en lo aprendido, reflexionando de lo aprendido, por ejemplo: con una simulación o con autoevaluaciones.

Dimensión 4: Progreso de la información

¿Cómo avanza el estudiante en su aprendizaje?

Secuencial: por pasos.

Global: de manera integral.

Apéndice G. Resumen de Moodle



Moodle es un paquete de Software para la creación y gestión de cursos en instituciones educativas a través de internet o de una intranet. En la actualidad es usado por gran número de instituciones educativas alrededor del mundo. La característica online ha permitido ese crecimiento.

Características

- Es a través de internet (facilita la comunicación a distancia).
- Gratuito (es de software libre).
- Fácil de instalar y actualizar.
- Soporte Técnico Comunitario.
- Uso intuitivo (tanto alumnos como profesores).
- Flexible y personalizable.

Actividades

- Foros y Chats.
- Mensajería Interna.
- Consultas y Encuestas.
- Calendario.

Actividades formativas

- Individuales: Lecciones, Tareas, Enlaces a Webs o documentos de cualquier formato.
- Colaborativas: Talleres, wikis, Diarios, Foros, Glosarios.
- De evaluación: Cuestionarios, Hot Potatoes, Exámenes auto evaluables.

Contenidos Integrables a Moodle

- Videos, Videotutoriales, Audios, Mp3, Presentaciones, Simulaciones.

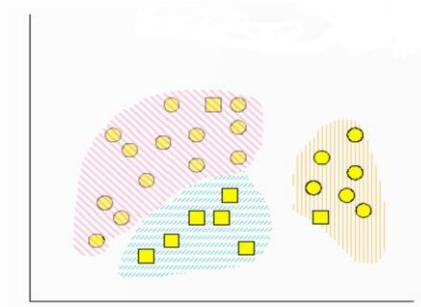
Apéndice H. Agrupamiento K-Means.

K-Means es un método de **agrupamiento**, que tiene como objetivo la partición de un conjunto de n observaciones en k grupos, en el que cada observación pertenece al grupo cuyo valor medio es más cercano. Es un método frecuentemente utilizado en minería de datos.

Dados unos datos sin etiquetar, el objetivo es encontrar grupos naturales

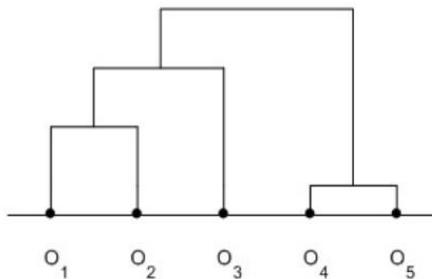
Tipos de Clustering

Particional: Partición de los objetos en grupos o clusters. Todos los objetos pertenecen a alguno de los k clusters, los cuales son disjuntos. El problema en este caso es la elección de k .



Particional

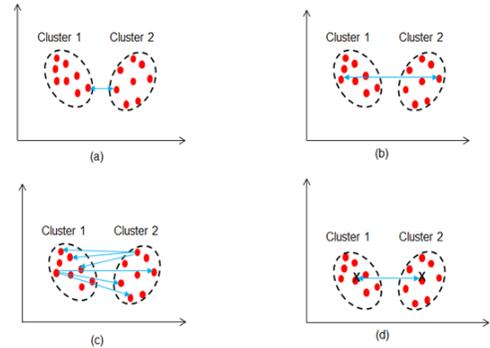
Jerárquico: Crear un dendograma, es decir, crear un conjunto de agrupaciones anidadas hasta construir un árbol jerárquico.



Jerárquico

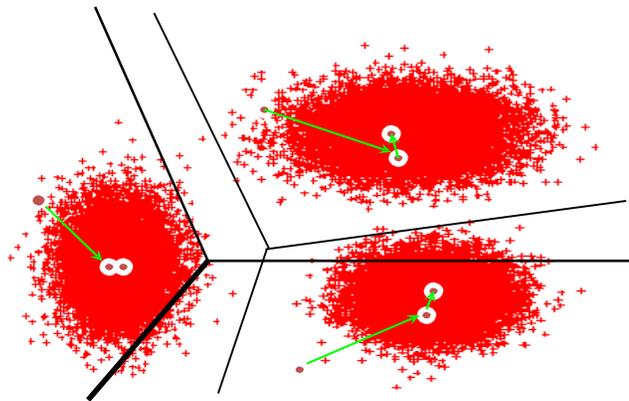
Tipos de Clustering basados en distancias

- (a) Distancia mínima.
- (b) Distancia máxima.
- (c) Distancia de promedio del grupo.
- (d) Distancia con respecto al centroide.



Algoritmo K-Means

1. Sitúa K puntos en el espacio en el que "viven" los objetos que se quieren clasificar. Estos puntos representan los centroides iniciales de los grupos.
2. Asigna cada objeto al grupo que tiene el centroide más cercano.
3. Tras haber asignado todos los objetos, recalcula las posiciones de los K centroides.
4. Repite los pasos 2 y 3 hasta que los centroides se mantengan estables. Esto produce una clasificación de los objetos en grupos que permite dar una métrica entre ellos.



Clusterización Algoritmo K-Means