

Metacognición e Inteligencia Artificial

Más allá del paralelismo de funcionamiento

Programa de Doctorado en Tecnologías
Industriales e Ingeniería Civil

Por María Consuelo Sáiz Manzanares



Universidad de Burgos

Memoria presentada para optar al Título de doctora
por la Universidad de Burgos dentro del Programa de Doctorado
en Tecnologías Industriales e Ingeniería Civil

DIRECTORES:

DR. CÉSAR IGNACIO GARCÍA OSORIO

DR. JOSÉ FRANCISCO DIEZ PASTOR

ÁREA DE LENGUAJES Y SISTEMAS INFORMÁTICOS

DEPARTAMENTO DE INGENIERÍA CIVIL

Burgos octubre de 2019

A mi hija Rut por aceptar mis sueños

A mi padre, donde esté, por enseñarme a amar la precisión en el trabajo

ÍNDICE

AGRADECIMIENTOS.....	1
RESUMEN.....	3
ABSTRACT	4
1. INTRODUCCIÓN.....	7
1.1. Metacognición, autorregulación y mejora del aprendizaje.....	14
1.2. Metacognición y feedback	16
1.3. Metacognición, autorregulación y feedback en entornos colaborativos.....	19
1.3.1. Computer Supported Collaborative Learning (CSCL)	20
1.3.2. Computer Based Learning Environment (CBLE).....	21
1.3.3. Diferencias y similitudes entre CBLE y CSCL.....	23
1.4. Metacognición e Inteligencia Artificial	23
1.5. Aprendizaje Basado en Proyectos: desarrollo en entornos virtuales	25
1.6. El proceso de enseñanza-aprendizaje en LMS.....	28
1.6.1. Logs, Analíticas de aprendizaje (LA) y Minería de datos educativos (EMD)...	30
1.7. Personalización del aprendizaje y e-orientación.....	36
1.7.1. Feedback y tutoría programada en la Universidad.....	37
1.7.2. Detección del alumno en riesgo de fracaso a través de las LMS	38
1.7.3. La personalización de aprendizaje	39
1.7.4. Técnicas de EMD y LA aplicadas al estudio del desarrollo de la asignatura «Estimulación Temprana».....	50
1.7.5. Recursos hipermedia en la personalización del aprendizaje	57
1.8. Técnicas Data Mining (DM) y de Pattern Mining (PM) en el entorno Educativo ...	60
1.9. Técnicas de Machine Learning utilizadas.....	71
1.9.1. Técnicas de aprendizaje supervisado de clasificación	71
1.9.2. Técnicas de aprendizaje supervisado de Regresión	75
1.9.3. Técnicas de aprendizaje no supervisado	84
2. PUBLICACIONES Y ACCIONES DE DIFUSIÓN DE LA INVESTIGACIÓN	97
2.1. Publicaciones principales.....	97
2.2. Publicaciones secundarias.....	105
2.3. Otras acciones de difusión de la investigación	109

2.4. Relación de las publicaciones y las técnicas de estadísticas y de minería de datos implementadas	119
2.4.1. Publicaciones principales.....	119
2.4.2. Publicaciones secundarias.....	120
APÉNDICE I. Publicaciones principales	123
APÉNDICE II. CUADERNO DE CAMPO (estudios piloto).....	125
3.1. Objetivos.....	125
3.2. Preguntas de Investigación	126
3.3. Método.....	127
3.3.1. Diseños.....	127
3.3.2. Participantes	129
3.3.3. Instrumentos.....	130
3.3.4. Procedimientos.....	144
3.4. Resultados.....	148
3.5. Conclusiones.....	177
3.6. Líneas futuras de investigación	182
3.7. Referencias Bibliográficas.....	185
3.8. Relación entre los estudios piloto y las publicaciones y acciones de difusión.....	205
RELACIÓN DE FIGURAS Y DE TABLAS	207
Relación de Figuras y de Tablas. Punto 1. Introducción	207
Relación de Figuras y de Tablas. Cuaderno de campo.....	208
GLOSARIO DE TÉRMINOS	211
RELACIÓN DE ACRÓNIMOS	217

AGRADECIMIENTOS

En primer lugar, quisiera agradecer al Área de Lenguajes y Sistemas Informáticos su apoyo, cariño y aceptación, pese a no ser informática, a lo largo de estos cuatro años de colaboración, especialmente a: Raúl Marticorena por iniciarme en el fantástico mundo de la Enseñanza Virtual, lo que ha cambiado mi forma de ver y de practicar la enseñanza. Durante estos años siempre ha tenido tiempo para enseñarme y para acompañarme en mi aprendizaje con una paciencia infinita; Álgvar Arnaiz Gonzalez, por orientar hacia el conocimiento tecnológico nuestro Grupo de Innovación Docente *Blended Learning* en Ciencias de la Salud, por tener siempre una sonrisa y tiempo para enseñarme y corregir mis inicios en las técnicas de aprendizaje automático en WEKA; Carlos Pardo, por apoyar siempre mis proyectos tanto dentro como fuera del trabajo de esta tesis, y por supuesto a mis directores de tesis César García Osorio, por enseñarme con mucha paciencia Python y técnicas de aprendizaje automático y por orientar de forma precisa mi aprendizaje en Técnicas de Minería de Datos y José Francisco Díez Pastor, por aceptar desde el inicio mi interés en conocer más profundamente las técnicas de Inteligencia Artificial.

A mi hija Rut, por acompañarme siempre con una sonrisa en todos mis proyectos, aunque nos hayan quitado tiempo de interacción, por aceptarme como soy y comprender que seguir aprendiendo es para mí una necesidad para seguir viviendo.

A la memoria de mi padre, que desde donde esté sé que verá con orgullo como sus enseñanzas me han acompañado y acompañan en mi camino.

Al director de mi primera tesis doctoral, José María Román, por orientar y apoyar mi trabajo en el campo de la metacognición y su aplicación a la enseñanza a lo largo de estos 25 años. Esta nueva tesis no es más que la continuación y actualización de la primera al campo de las nuevas tecnologías y técnicas de análisis de los datos en el ámbito de la minería de datos.

A todos los que me acompañáis en el día a día con vuestro cariño y apoyo.

RESUMEN

El objeto de esta tesis ha sido el de aplicar conocimientos de la Inteligencia Artificial al campo de la Psicología Cognitiva y Metacognitiva. En concreto, este trabajo se ha centrado en el uso de técnicas de aprendizaje automático (*Machine Learning*), tanto respecto del uso de técnicas de aprendizaje supervisado (*supervised*) como de las técnicas de aprendizaje no supervisado (*unsupervised*). Ambas se han aplicado para favorecer el aprendizaje personalizado en los estudiantes y, en último término mejorar su rendimiento y satisfacción. En la actualidad el aprendizaje, especialmente en entornos de Educación Superior, se está desarrollando a través de plataformas interactivas, *Learning Management System* (LMS), como por ejemplo Moodle. Estos entornos ofrecen muchas posibilidades de registro de la actividad (logs o unidades de registro) y de extracción de dicha información, ya que estos registros se almacenan en bases de datos y estas pueden ser descargadas de forma sencilla en ficheros con distintos formatos CSV, XLS, JSON, etc. No obstante, los LMS no siempre contienen herramientas para el análisis de la información. Habitualmente estos entornos ofrecen algunos procedimientos de *Learning Analytics* que sólo permiten análisis de datos sencillos. Es por lo que, si se quieren realizar estudios más precisos, se deben aplicar herramientas más complejas que incluyan técnicas estadísticas y de minería de datos más sofisticadas. Especialmente, respecto de las segundas, se incluyen las técnicas de clasificación, de regresión y de clustering. Dichas herramientas pueden ser SPSS, R, WEKA o bibliotecas de Python como Pandas y *Scikit-Learn*.

La tesis se divide en dos capítulos y dos apéndices, en el Capítulo 1 se presenta un acercamiento a los conceptos de aprendizaje, metacognición, autorregulación y feedback. Seguidamente, se relacionan dichos conceptos con el aprendizaje en entornos virtuales. Más adelante, se explica la relación entre los conceptos de Metacognición, Inteligencia Artificial, Aprendizaje Basado en Proyectos en entornos virtuales, proceso de enseñanza aprendizaje en entornos virtuales, procesos y procedimientos de análisis de datos, aplicación de técnicas de minería de datos y finalmente la aplicación de todo ello en la personalización del aprendizaje y en la propuestas de módulos de e-orientación. En el Capítulo 2, se recogen las aportaciones de esta tesis: artículos, libros, capítulos de libro y acciones de difusión de la investigación como simposios y conferencias invitadas y acciones formativas.

Asimismo, en el Apéndice I, se presentan las publicaciones principales de la tesis y en el Apéndice II, se incluye un cuaderno de campo en el que se exponen todos los trabajos previos a los artículos publicados y que sirvieron de base para estos. En concreto tres experimentos piloto que se plantean con la siguiente estructura metodológica: objetivos y las preguntas de investigación, método, instrumentos y procedimientos, resultados, conclusiones y las líneas futuras de intervención.

Palabras clave: metacognición, técnicas de minería de datos educativas, inteligencia artificial, aprendizaje basado en proyectos, e-orientación.

ABSTRACT

The aim of this thesis has been to apply knowledge of artificial intelligence to the field of cognitive and metacognitive psychology. Specifically, this work has focused on the use of automatic learning techniques (Machine Learning), supervised and unsupervised, to analyze the effect on the satisfaction and performance of students of different actions that sought to provide them with a personalized learning experience. Currently, learning, especially in Higher Education environments, is being developed through interactive platforms, Learning Management System (LMS), such as Moodle. These environments offer many possibilities for registering the student activity (logs or registration units) and extracting this information, as these records are stored in databases and these can be downloaded easily in files with different formats CSV, XLS, JSON, and so on. However, LMS do not always contain tools for information analysis. Usually these environments offer Learning Analytics procedures that only include simple data analysis techniques. This is why, if you want to perform more complex studies, you must apply more complex tools that include more sophisticated statistical and data mining techniques. The latter include classification, regression and clustering techniques. These techniques are provided by tools like SPSS, R, WEKA, or Python libraries such as Pandas and Scikit-Learn.

The thesis is divided into two chapters and two appendixes, Chapter 1 presents an approach to the concepts of learning, metacognition, self-regulation and feedback. These concepts are then related to learning in virtual environments. Later, a relationship is made between the concept of metacognition and that of Artificial Intelligence, Project Based Learning Methodology in virtual environments, process of teaching learning in virtual environments and processes and procedures of data analysis and application of Data Mining

Techniques and finally the application of all this in the personalization of learning and in the proposals of e-orientation modules. Chapter 2 contains the contributions of this thesis: articles, books, book chapters and research dissemination actions such as symposia and invited conferences and training actions.

Also, Appendix I, the principal contributions of this thesis are presented, and Appendix II includes a field notebook in which all the works prior to the published articles are presented and which served as a basis for them. Specifically, three pilot experiments are described with the following methodological structure: objectives and research questions, method, instruments and procedures, results, conclusions and future lines of intervention.

Keywords: metacognition, Educational Data Mining, Artificial Intelligence, Project-based learning, e-orientation

RESUMEN

El objeto de esta tesis ha sido el de aplicar conocimientos de la inteligencia artificial al campo de la Psicología cognitiva y metacognitiva. En concreto, este trabajo se ha centrado en el uso de técnicas de aprendizaje automático (*Machine Learning*), tanto respecto del uso de técnicas de aprendizaje supervisado (*supervised*) como de las técnicas de aprendizaje no supervisado (*unsupervised*). Ambas se han aplicado para favorecer el aprendizaje personalizado en los estudiantes y, en último término mejorar su rendimiento y satisfacción. En la actualidad el aprendizaje, especialmente en entornos de Educación Superior, se está desarrollando a través de plataformas interactivas, *Learning Management System* (LMS), como por ejemplo Moodle. Estos entornos ofrecen muchas posibilidades de registro de la actividad (*logs* o unidades de registro) y de extracción de dicha información, ya que estos registros se almacenan en bases de datos y estas pueden ser descargadas de forma sencilla en ficheros con distintos formatos CSV, XLS, JSON, etc. No obstante, los LMS no siempre contienen herramientas para el análisis de la información. Habitualmente estos entornos ofrecen procedimientos de *Learning Analytics* que incluyen técnicas de análisis de datos sencillas. Es por lo que, si se quieren realizar estudios más precisos se deben aplicar herramientas más complejas que incluyan técnicas estadísticas y de minería de datos más sofisticadas. Especialmente, respecto de las segundas se incluyen las técnicas de clasificación, de regresión y de *clustering*. Dichas herramientas pueden ser SPSS, R, WEKA o bibliotecas de Python como Pandas y *Scikit-Learn*.

La tesis se divide en tres capítulos, en el Capítulo 1 se presenta un acercamiento a los conceptos de aprendizaje, metacognición, autorregulación y *feedback*. Seguidamente, se relacionan dichos conceptos con el aprendizaje en entornos virtuales. Más adelante, se hace una relación entre el concepto de Metacognición y el de Inteligencia Artificial, Metodología de Aprendizaje Basado en Proyectos en entornos virtuales, proceso de enseñanza aprendizaje en entornos virtuales y procesos y procedimientos de análisis de datos y aplicación de Técnicas de Minería de Datos y finalmente la aplicación de todo ello en la personalización del aprendizaje y en la propuestas de módulos de e-orientación.

En el Capítulo 2, se recogen las aportaciones de esta tesis: artículos, libros, capítulos de libro y acciones de difusión de la investigación como simposios y conferencias invitadas y acciones formativas.

Asimismo, en el Apéndice II, se incluye un cuaderno de campo en el que se exponen todos los trabajos previos a los artículos publicados y que sirvieron de base para ello. En concreto tres experimentos piloto que se plantean con la siguiente estructura metodológica: objetivos y las preguntas de investigación, método, instrumentos y procedimientos, resultados, conclusiones y las líneas futuras de intervención.

Palabras clave: metacognición, técnicas de minería de datos educativas, inteligencia artificial, aprendizaje basado en proyectos, e-orientación

ABSTRACT

The object of this thesis has been to apply knowledge of artificial intelligence to the field of cognitive and metacognitive psychology. Specifically, this work has focused on the use of automatic learning techniques (Machine Learning), both regarding the use of supervised learning techniques and unsupervised learning techniques. Both have been applied to encourage personalized learning in students and ultimately improve their performance and satisfaction. Currently learning, especially in Higher Education environments, is being developed through interactive platforms, Learning Management System (LMS), such as Moodle. These environments offer many possibilities for registering the activity (logs or registration units) and extracting this information, as these records are stored in databases and these can be downloaded easily in files with different formats CSV, XLS, JSON, and so on. However, LMS do not always contain tools for information analysis. Usually these environments offer Learning Analytics procedures that include simple data analysis techniques. This is why, if you want to perform more complex studies, you must apply more complex tools that include more sophisticated statistical and data mining techniques. The latter include classification, regression and clustering techniques. These tools can be SPSS, R, WEKA or Python libraries such as Pandas and Scikit-Learn.

The thesis is divided into three chapters, Chapter 1 presents an approach to the concepts of learning, metacognition, self-regulation and feedback. These concepts are then related to learning in virtual environments. Later, a relationship is made between the concept of metacognition and that of Artificial Intelligence, Project Based Learning Methodology in virtual environments, process of teaching learning in virtual environments and processes and procedures of data analysis and application of Data Mining Techniques and finally the

application of all this in the personalization of learning and in the proposals of e-orientation modules.

Chapter 2 contains the contributions of this thesis: articles, books, book chapters and research dissemination actions such as symposia and invited conferences and training actions.

Also, Appendix II includes a field notebook in which all the works prior to the published articles are presented and which served as a basis for this. Specifically, three pilot experiments are proposed with the following methodological structure: objectives and research questions, method, instruments and procedures, results, conclusions and future lines of intervention.

Keywords: metacognition, Educational Data Mining, Artificial Intelligence, Project-based learning, e-orientation

1. INTRODUCCIÓN

Este trabajo parte de los estudios de investigación realizados en el campo de la Ciencia Cognitiva y Metacognitiva. El concepto de metacognición hace referencia a la reflexión consciente que un sujeto hace sobre su propio conocimiento. Dicha reflexión implica la toma de conciencia de qué tipo de problema hay que resolver, para qué resolverlo y cómo se va a resolver. A su vez, el espacio metacognitivo está conformado por una serie de factores (Flavell, 1979):

1. Fines o metas cognitivas, relacionadas con las tareas a resolver.
2. Experiencias metacognitivas, información que los sujetos tienen acerca de la utilización de sus habilidades cognitivas en tareas semejantes.
3. Acciones cognitivas, estrategias que un sujeto emplea en la resolución de problemas.
4. Conocimiento metacognitivo, conocimiento acerca de las variables persona, tarea y estrategias.

Un término directamente relacionado con el concepto de metacognición, es el de autorregulación (Brown, 1987). La autorregulación va a permitir al sujeto la elaboración y el auto-análisis de los heurísticos de resolución de problemas. Dicho concepto se relaciona con los procesos de planificación, *feedback* y evaluación en el proceso de aprendizaje (Brown, 1987; Brown y DeLoache, 1978; Sáiz y Román, 2011), la *self-efficacy* (Bandura, 1977) y el desarrollo de la motivación en el proceso de aprendizaje (Zimmerman, 2008). Existen diferentes planteamientos teóricos sobre la relación entre metacognición y autorregulación. Si bien, actualmente los estudios se decantan por entender que la autorregulación está supraordenada a la metacognición en las respuestas de aprendizaje (*learning outcomes*) (citado por Veenman, 2011 p. 197). La autorregulación se relaciona con el conocimiento metacognitivo y con la autorregulación metacognitiva, esta última hace referencia a las habilidades de autorregulación, (*self-regulatory skills*) de los estudiantes, necesarias para activar el auto-control del aprendizaje de forma coordinada (Greene y Azevedo 2009; Meijer, Veenman, y Van Hout-Wolters, 2006). No obstante, la regulación metacognitiva parece no ser suficiente para producir una regulación adecuada del propio aprendizaje. Por ello, para conseguir un pleno desarrollo del proceso de autorregulación se precisará de un diseño, implementación y evaluación de las iniciativas de metacognición autorregulada (De Backer, Van Keer, y Valcke, 2015). Así pues, esta es entendida como la

auto-conciencia del aprendiz sobre su propio aprendizaje y sobre el control respecto de las actividades cognitivas que realiza durante la resolución de tareas (ver Figura 1).

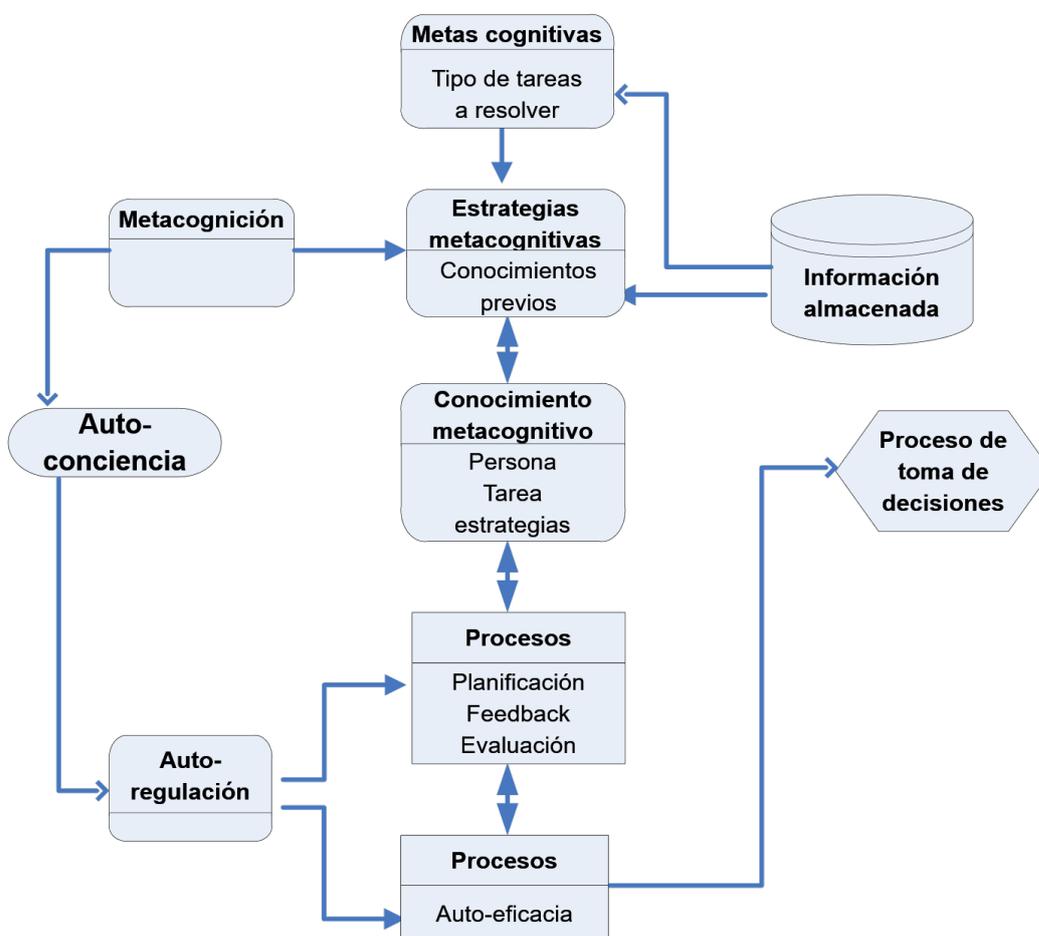


Figura 1

Relación entre los conceptos de metacognición y de autorregulación

El primer componente es el conocimiento metacognitivo (*metacognitive knowledge*) y el segundo la autorregulación metacognitiva (*metacognitive regulation*) (De Backer et al., 2015). En este último concepto, se pueden diferenciar los siguientes componentes: Orientación (*Orientation*), implica ajustar los objetivos de la tarea y relacionarlos con los conocimientos previos; Planificación (*Planning*), incluye la selección y secuenciación de las estrategias de resolución del problema y la asignación de recursos y de planes de acción antes o durante la ejecución (Greene y Azevedo, 2009); Monitorización (*Monitoring*), contempla el control durante la ejecución e implica la identificación de las tareas de ejecución, la corrección en su caso de las acciones (Greene y Azevedo 2009; Meijer et al., 2006; Moos, y Azevedo 2009) este constructo está directamente relacionado con el progreso

en la profundidad en el aprendizaje, y Evaluación (*Evaluation*), implica la comprobación de las metas (Veenman y Spaans, 2005) este concepto está relacionado con las respuestas (*learning outcomes*) y los procesos de aprendizaje (*learning process*) (Meijer et al. 2006) e implica el éxito final en las respuestas de aprendizaje (ver Figura 2). El buen desarrollo de todas estas estrategias es la clave para un efectivo uso de las estrategias de regulación. Una de las estrategias que el docente puede utilizar para potenciar estas habilidades es la regulación escalonada (*scaffolding*) de las metas de aprendizaje a través de aproximaciones sucesivas a la meta (Feng, 2014).

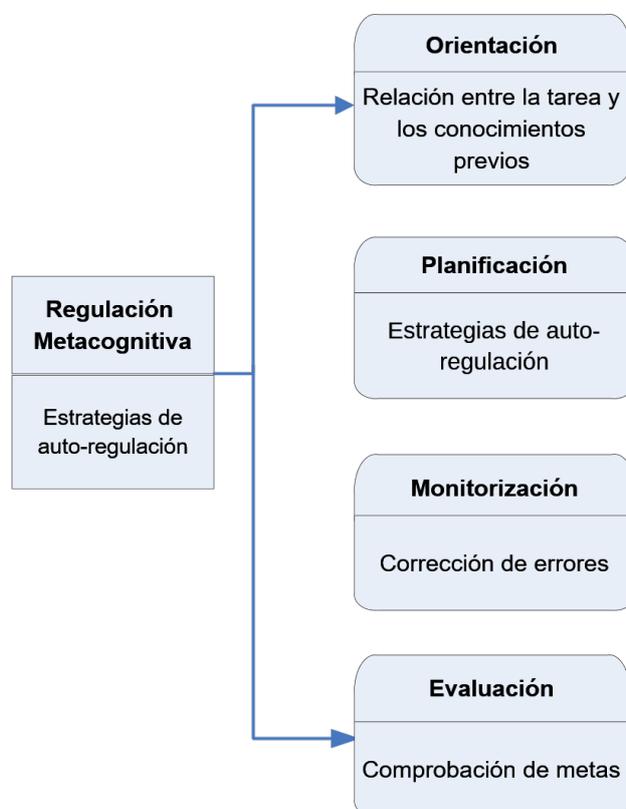


Figura 2

Regulación metacognitiva y los componentes metacognitivos

Dentro del término metacognición se pueden diferenciar siguientes componentes (Veenman, 2011):

1. Conocimiento de la cognición (*Knowledge of cognition*), incluye el conocimiento declarativo (*declarative knowledge*) sobre el sistema cognitivo, el conocimiento

procedimental (*procedural knowledge*), que hace referencia a las estrategias cognitivas y metacognitivas de resolución. Estas dos estrategias guían y controlan el propio aprendizaje en la resolución de problemas (Veenman y Elshout, 1999). Por último, también incluye el conocimiento condicional (*conditional knowledge*) que hace referencia a cuándo y para qué utilizar determinadas estrategias metacognitivas en la resolución de una tarea.

2. Regulación de la cognición (*Cognition Regulation*), hace referencia a las estrategias metacognitivas necesarias para el control en el uso de las estrategias metacognitivas más adecuadas en la resolución de tareas, incluye las estrategias de planificación (*planning*), monitorización o supervisión (*monitoring*) y las de evaluación (*evaluation*).
3. Conocimiento metacognitivo (*Metacognitive knowledge*), se basa en la autoconciencia que un sujeto tiene de sus estados mentales, estrategias de aprendizaje y el uso de dichas estrategias en la resolución de problemas o en las actividades de aprendizaje (De Backer et al., 2015). El conocimiento metacognitivo interactúa con otros aspectos de la metacognición como son las experiencias metacognitivas (*metacognitive experiences*) y las estrategias metacognitivas (*metacognitive skills*) (Efklides, 2009).
4. Experiencias metacognitivas (*Metacognitive experiences*), hace referencia a la experiencia del sujeto en la resolución de tareas semejantes. Este concepto se interrelaciona con aspectos de la persona y de la tarea e incluye los sentimientos, juicios y tareas específicas del conocimiento (*metacognitive feelings, metacognitive judgments y task-specific knowledge*). Todos estos conceptos ayudan a calibrar el esfuerzo y el tiempo de resolución de una tarea (Flavell, 1979) y hacen referencia a las estrategias metacognitivas de orientación (*Orientation*).

De otro lado, Wäschle, Allagaier, Lachner, Siegfried, y Nückles (2013) identificaron los siguientes componentes dentro del concepto de metacognición:

1. Establecimiento de metas (*Goal-setting*), definen los pasos en la resolución de un problema.
2. Objetivos de aprendizaje auto-definidos (*Self-set learning goals*), guían el aprendizaje y regulan qué pasos se seguirán y por qué.

3. Estrategias de aprendizaje (*Learning strategies*), serían las estrategias de adquisición, codificación, recuperación, metacognitivas y de apoyo al procesamiento de la información.
4. Tipo de tareas (*Type of tasks*), en ellas se diferencia entre el logro del alumno de un aprendizaje profundo o superficial, así como del uso de las estrategias de resolución implicadas en cada uno de ellos.
5. Auto-reflexión sobre el logro de las metas (*Self-reflection of goal achievement*), ayudan a los estudiantes a reflexionar sobre las producciones de aprendizaje.

Para entender el proceso autorregulatorio, Zimmerman y Moylan (2009) proponen un modelo de fase de retroalimentación basado en el bucle de autorregulación (*Phase Model of Self-regulatory Feedback loop*) en el que se diferencian tres fases interconectadas (ver Figura 3)

1. Fase de previsión (*Forethought phase*), incluye *task analysis processes* que incluyen estrategias de planificación y de auto-motivación.
2. Fase de desarrollo (*Performance phase*), incluye los métodos de auto-control (*self-control*) y auto-observación (*self-observation*).
3. Fase de auto-reflexión (*Self-Reflection phase*), se compone de auto-juicios (*self-judgements*) que incluyen la auto-evaluación (*self-evaluation*), estas estrategias pueden ser potenciadas por el profesor o por los iguales, y auto-reacciones (*self-reactions*) que están relacionadas con la auto-satisfacción y la motivación.

Todos ellos son poderosos predictores del aprendizaje eficaz (Zimmerman y Moylan, 2009).

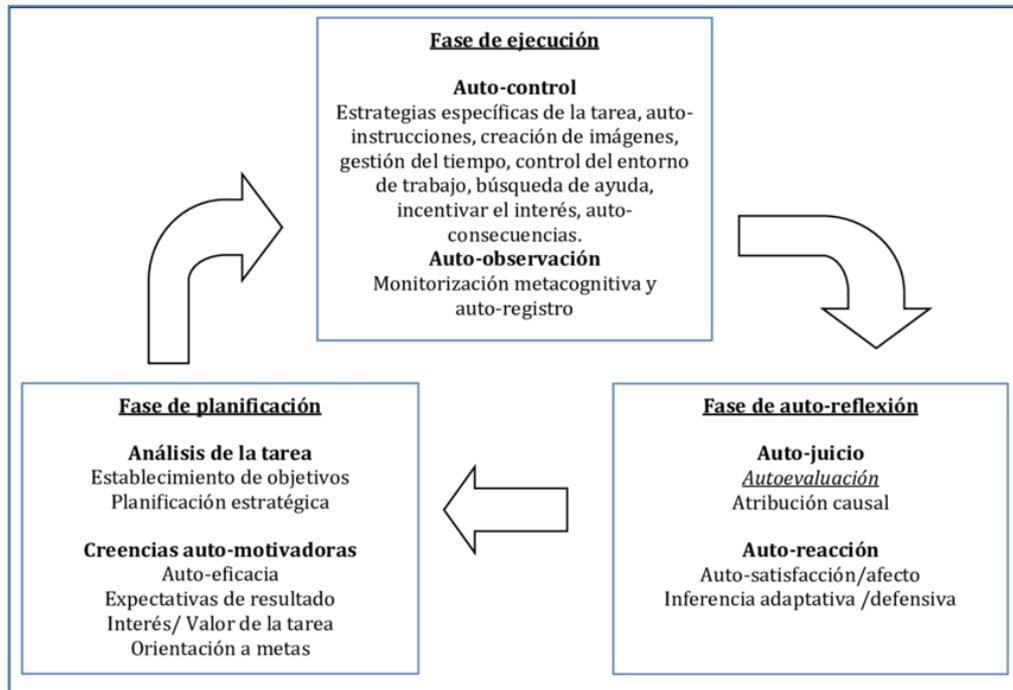


Figura 3

Modelo de Zimmerman y Moylan (2009)

En la Figura 4 se pueden observar la relación entre los conceptos de conocimiento de la cognición, regulación de la cognición y conocimiento metacognitivo.

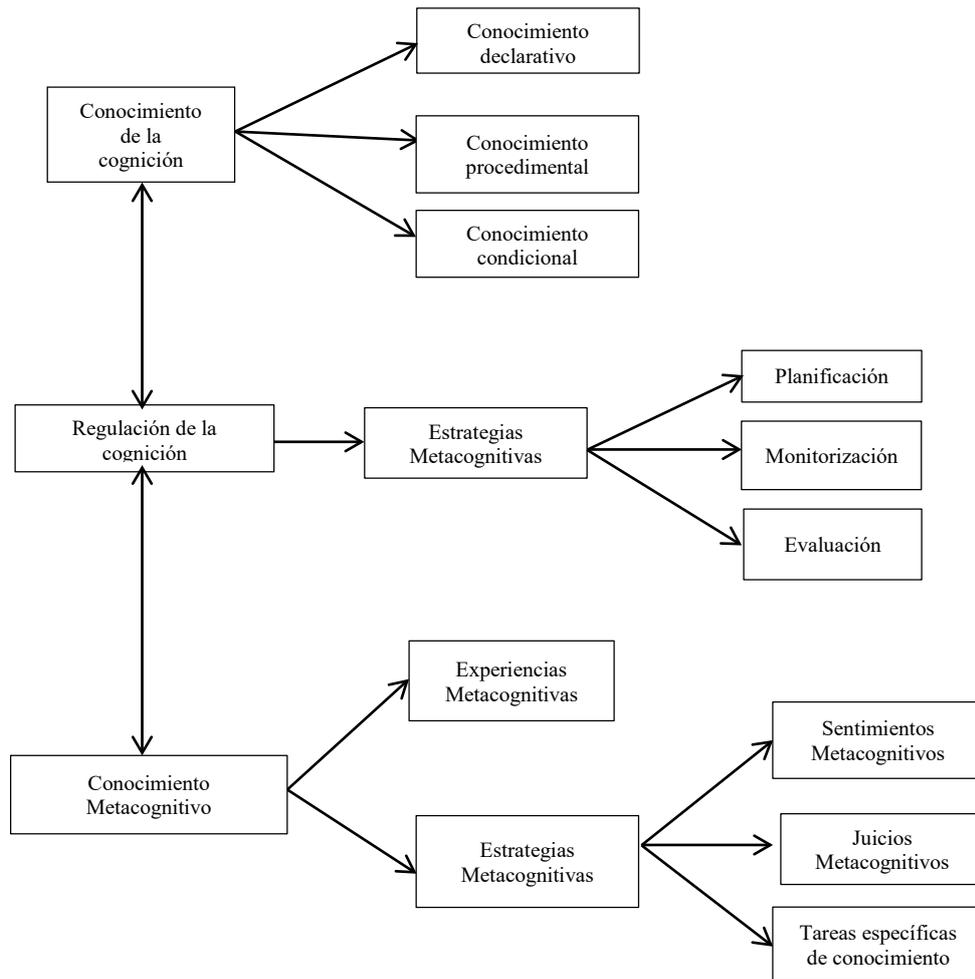


Figura 4

Relación entre los conceptos de conocimiento de la cognición, regulación de la cognición y conocimiento metacognitivo

Además de todo lo señalado anteriormente, para desarrollar aprendizajes exitosos hay que tener en cuenta las experiencias metacognitivas que el sujeto ha tenido respecto de la tarea objeto de aprendizaje (Efklides, 2009). Este es el motivo por el que los aprendices nóveles durante la resolución de tareas no habituales pueden, si no se les apoya, desarrollar aprendizajes erróneos o superficiales. Para evitarlo, el diseño de la instrucción de nuevos aprendizajes tiene que estar estructurado de la siguiente forma:

- Inicio del desarrollo de la tarea: se analiza la tarea, implica la activación de los conocimientos previos, el establecimiento de las metas, la planificación de resolución y las estrategias de regulación de los procesos cognitivos.

- Durante el desarrollo de la tarea: se sigue de forma sistemática el plan que se ha establecido observando (*monitoring*) y supervisando (*checking*) el proceso. Estas actividades guían y controlan la ejecución de la tarea.
- Al final del desarrollo de la tarea: se evalúa y se reflexiona, tanto sobre el proceso de resolución, como sobre los resultados finales. Fruto de ello, se obtiene un aprendizaje acerca del procedimiento para resolver tareas semejantes.

1.1. Metacognición, autorregulación y mejora del aprendizaje

Nelson y Narens (1990) diferenciaron entre el *object level* y el *meta level*. El primer nivel (*object level*) ejecuta los procesos cognitivos de nivel más bajo (procesos de memoria, de lectura y de resolución de problemas) y el segundo, *meta level*, hace referencia a los procesos metacognitivos (comparación, razonamiento e inferencias). Dichos procesos implican estrategias de planificación y evaluación que permiten el análisis sobre el *object level*. La monitorización (supervisión) y el análisis del error, serían procesos *bottom-up* situados en el *object level*. Asimismo, los procesos de elaboración se sitúan en los procesos *top-down* ubicados en el *meta-level*. En este último, se incluye las estrategias metacognitivas que se relacionan con las estrategias de auto-instrucción que ayudan a regular los procesos cognitivos. Dichas estrategias de auto-instrucción se activan en el desarrollo de la resolución de problemas y contienen el conocimiento condicional sobre dónde aplicar de forma particular las habilidades metacognitivas y las instrucciones procedimentales para saber cómo implementar las habilidades del *object level*. En la fase cognitiva se hace referencia al conocimiento declarativo de condición y de acción que permite una descripción verbal de la estrategia respondiendo a las cuestiones: qué hacer (*What to do*), cuándo, porqué y cómo (*When, Why* y *How*) (Veenman et al., 2006 p. 9). Asimismo, el conocimiento metacognitivo, y en particular conocimiento condicional, incluye la regla *if-then*. En las fases iniciales de la resolución de problemas las estrategias metacognitivas precisan desarrollarse paso a paso con el fin de reducir los errores. Posteriormente, en las fases subsiguientes del proceso de resolución, el conocimiento procedimental (procedural) tendrá que ser graduado para mejorar la fluidez y permitir una detección temprana de los errores y la corrección autónoma de los mismos (Veenman, 2015). El esquema de Nelson y Narens (1990), se puede consultar en el Figura 5.

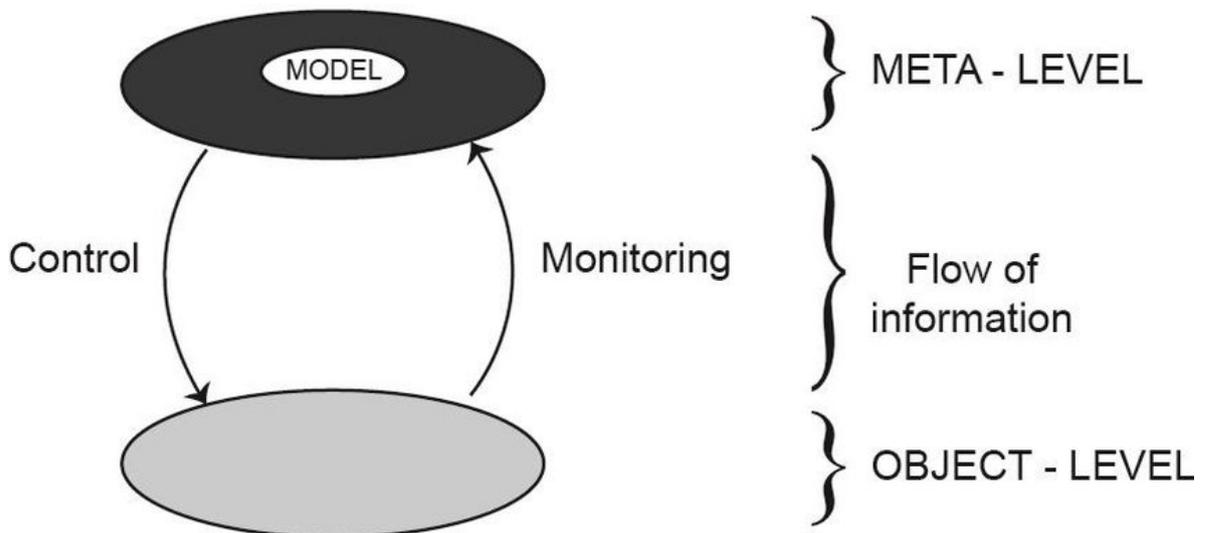


Figura 5

Modelo de Nelson y Narens (1990)

La investigación en Psicología Cognitiva a lo largo de 3 décadas revela que el uso de las estrategias metacognitivas y el empleo de las estrategias de autorregulación facilitan en los estudiantes el desarrollo de aprendizajes seguros, profundos y eficaces (Azevedo, 2014; Malmberg, Järvelä, Järvenoja, y Panadero, 2015). Es por lo que el entrenamiento durante la instrucción de dichas estrategias va a mejorar en los estudiantes la inhibición de los procesos de ansiedad. Lo que previsiblemente incrementará el desarrollo de la motivación y de los aprendizajes exitosos (Forstmann y Wagenmakers, 2015; Kostons, Van Gog, y Paas, 2012; Schraw y Gutiérrez, 2015). En la Tabla 1 pueden consultarse las relaciones entre las estrategias de aprendizaje en los distintos momentos del proceso de resolución de problemas.

Tabla 1

Relación entre los conceptos de conocimiento de la cognición, regulación de la cognición y conocimiento metacognitivo

Estrategias	Antes del aprendizaje	Durante el aprendizaje	Después del aprendizaje
Orientación	<ul style="list-style-type: none"> Activar los conocimientos previos. Identificar los problemas. Leer las instrucciones. Definir las metas de aprendizaje. Realizar predicciones del posible proceso de resolución. 	<ul style="list-style-type: none"> Relacionar lo nuevo con lo que ya se sabía. Utilizar mapas conceptuales y diagramas de flujo que ayuden a organizar las ideas. 	
Planificación	<ul style="list-style-type: none"> Realizar un plan de resolución. Realizarse auto-preguntas. Buscar espacio adecuados de aprendizaje. Elegir las estrategias adecuadas para cada uno de los objetivos del plan de acción basadas en conocimiento condicional. Efectuar una revisión final. 	<ul style="list-style-type: none"> Reevaluar los planes. 	
Evaluación y auto-supervisión	<ul style="list-style-type: none"> Detectar errores Verificar respuestas y soluciones. Auto-corrección 	<ul style="list-style-type: none"> Organizar la información. Hacer inferencias. Utilizar protocolos de <i>think aloud</i> y reflexionar sobre su contenido. Construir un <i>metalevel</i> en la propia estructuración mental. Revisar la comprensión. Evaluar el logro de las metas de aprendizaje. Buscar ayuda en aspectos en los que se tienen dudas. Cambiar estrategias. Reevaluar el trabajo. 	<ul style="list-style-type: none"> Comparar los resultados de aprendizaje con la planificación previa y con las evidencias del desarrollo del aprendizaje. Comparación entre los objetivos y los resultados de aprendizaje.
Reelaboración y reflexión	<ul style="list-style-type: none"> Realizar un resumen del proceso de resolución. Sacar conclusiones de lo aprendido y del proceso de aprendizaje. Detectar los aspectos generales que ayuden a la generalización de lo aprendido para la resolución de tareas semejantes. 		

1.2. Metacognición y feedback

Como se ha señalado anteriormente, la metacognición puede relacionarse con dos procesos, uno de monitorización y otro de control. Asimismo, algunos autores (Cox, 2005; Cox y Raja, 2012) la relacionan con los procesos de meta-razonamiento. No obstante, al

inicio de un nuevo aprendizaje, el aprendiz novel, en un primer momento, tiene que integrar los conocimientos nuevos con los que ya posee, es decir, tiene que efectuar cognitivamente un proceso de asimilación y acomodación de la nueva información para poder reestablecer un nuevo estado de equilibración (Piaget, 1975). Por ello, los niveles iniciales de la utilización de las habilidades metacognitivas no se desarrollará de forma completa. Además, hay que tener en cuenta que si en este complejo proceso no se aplica un *feedback* externo (docente, grupo de iguales) en los aprendices puede producirse un abandono en la resolución de las tareas. Asimismo, hay que considerar que la efectividad del *feedback* depende de su estructura y este tiene que guiar al aprendiz en los siguientes aspectos:

1. El acercamiento a los nuevos conceptos.
2. La resolución de problemas iniciales relacionados con los nuevos conceptos.
3. La estructuración cognitiva que permita la monitorización consciente de los procesos de resolución de problemas.
4. La corrección de los posibles errores en el algoritmo de resolución de problemas.
5. El uso de las estrategias de control de los procesos de resolución de problemas.
6. La adquisición de estrategias de meta-razonamiento que faciliten la generalización y la transferencia del aprendizaje.

La metacognición pues permite al aprendiz lograr un *insight* dentro del propio pensamiento, lo que facilita el desarrollo del aprendizaje autónomo y la motivación hacia la ejecución de la tarea (Schunk, Pintrich, y Meece, 2007; Butler et al., 2012; Aragão y Almeida, 2017).

El *feedback*, pues, se puede definir como la información que proviene de otros (docente o los iguales) sobre la ejecución de una tarea. El *feedback* ha mostrado tener una gran influencia sobre los procesos de aprendizaje (Hattie, 2008). El mayor porcentaje de los estudios sobre el *feedback* se han focalizado en su impacto en la corrección del error, en los cambios en las metas (*achievement change*), en la motivación y en la adquisición de meta o logros y en un menor porcentaje en el uso de las estrategias metacognitivas. Si bien, no todos los *feedback* tienen el mismo efecto (Hattie y Timperley, 2007). Así pues, la efectividad del *feedback* depende de su estructura y tiene que contemplar tres funciones en el aprendizaje (Harks, Rakoczy, Hattie, Besser, y Klieme, 2014):

- Cognitiva (orientado al logro): la información que se da al aprendiz debe estar lo suficientemente elaborada para ayudar al sujeto a detectar el error y a mejorar sus

resultados de aprendizaje. En síntesis, tiene que reducir la distancia entre la comprensión actual y las metas de aprendizaje marcadas.

- Motivacional (hacia el objeto de aprendizaje): la información que se otorga al aprendiz tiene que incrementar su motivación intrínseca para aprender la materia. Para que un *feedback* influya en la motivación tiene que:
 - tener criterios de referencia hacia el éxito en el aprendizaje.
 - referirse al aprendizaje y a los procesos de logro.
 - estabilizar las atribuciones en caso de fallos.
- Metacognitivo (auto-evaluación y elaboración de las respuestas): existe una relación entre el *feedback* y los procesos de autorregulación. Si se ayuda a los estudiantes a mejorar sus procesos de monitorización y la calibración sobre su proceso de aprendizaje previsiblemente mejorará el uso de las estrategias metacognitivas de auto-evaluación (Dinsmore y Parkinson, 2013; Hattie, 2013; Panadero y Alonso-Tapia, 2013).

Como ya se ha indicado el *feedback* posee una gran influencia en el proceso de aprendizaje (Hattie, 2008), aunque también tiene limitaciones (Sáiz y Montero, 2015). Dicho *feedback* está relacionado con la percepción que el sujeto tiene sobre su propia competencia de aprendizaje. Asimismo, su efectividad depende de la respuesta a las cuestiones:

- «¿A Dónde voy?»
- «¿Cómo voy a ir?»
- «¿Cuál es el paso siguiente?» (*where to go next?*).

Según Narciss (2008) el *feedback* con función de orientación proporciona al aprendiz refuerzos, facilita la resolución de las tareas e incrementa los niveles de motivación hacia el aprendizaje. Se pueden diferenciar tres niveles de *feedback* (Hattie y Gan, 2011):

1. *Feedback* sobre la tarea (*Feedback at task level*): da información sobre una respuesta correcta. Este nivel es importante cuando un aprendiz es novel ya que implica la relación entre el tipo de tarea y el proceso de aprendizaje de la misma (Malmberg, Järvelä, y Kirschner, 2014).
2. *Feedback* sobre el proceso (*Feedback at process level*): este proceso se relaciona con la auto-eficacia (*self-efficacy*) del sujeto en la resolución de la tarea.
3. *Feedback* a nivel de autorregulación (*Feedback at self-regulation level*): este nivel incluye estrategias de auto-evaluación (*self-evaluation*) y estrategias de búsqueda de

información (*feedback*). Todo ello implica la capacidad del aprendiz para crear sus propios procesos de *feedback* y de auto-evaluación (*self-assess*).

El nivel más alto de *feedback* es el tercero ya que comprende:

1. Componentes de orientación hacia una tarea concreta a resolver e información al estudiante sobre sus debilidades. También da respuesta a cuestiones como cuándo y porqué una estrategia es adecuada en el proceso de resolución.
2. Conocimiento sobre cómo continuar y sobre las estrategias de resolución de tareas.

El *feedback* (del docente o del grupo de iguales) sobre la tarea de aprendizaje se puede realizar *Face to Face* (F2F) o bien *online* (Hull y du Boulay, 2015).

En síntesis, la metacognición auto-regulada permite al aprendiz tener un *insight* dentro del propio proceso de resolución de tareas y fomenta el aprendizaje autónomo incrementando la motivación. Es decir, esta facilita el desarrollo de la auto-eficacia (*self-efficacy*) (Schunk et al., 2007). Asimismo, potencia la construcción de los aprendizajes, ya que orienta el uso de las representaciones mentales y de las relaciones cognitivas en la memoria, es decir de la recursividad del pensamiento (Breed, Mentz, y Van der Westhuizen, 2014; Falkner, Vivian, y Falkner, 2013).

Actualmente, la metacognición entendida desde su relación con la autorregulación, se considera un relevante predictor del aprendizaje (Primi, Ferrão, y Almeida, 2010). Las habilidades metacognitivas explican el 40 % de la varianza en los resultados de aprendizaje (Van der Stel y Veenman, 2008; 2010). Ambos conceptos, metacognición y autorregulación, están directamente vinculados al éxito en el aprendizaje (Meijer et al., 2006; Zimmerman y Schunk, 2011).

1.3. Metacognición, autorregulación y feedback en entornos colaborativos

En la última década se han desarrollado investigaciones relacionadas con la efectividad del aprendizaje en entornos virtuales colaborativos. Una parte de estos estudios se ha centrado en el aprendizaje de materias como la de programación orientada a objetos, ya que dicha materia permite la visualización en la resolución de tareas de los pasos tanto cognitivos como metacognitivos fundamentados desde la elaboración de un algoritmo (Sáiz

y Marticorena, 2016). Los resultados indican que la utilización de los entornos virtuales desde el aprendizaje colaborativo mejora los resultados de aprendizaje. Seguidamente se presentan dos tipos de entornos el *Computer Supported Collaborative Learning* (CSCL) y el *Computer Based learning environment* (CBLE). Ambos se han mostrado efectivos para el desarrollo de estrategias cognitivas y metacognitivas (Järvelä, Malmberg, y Koivuniemi, 2016). Una forma de comprobar su uso y efectividad es a través del diseño de rúbricas que pueden ser utilizadas tanto por el docente como por los alumnos (Boote, 2014; Sáiz y Bol, 2014).

1.3.1. Computer Supported Collaborative Learning (CSCL)

El CSCL tiene como objetivo mejorar la educación y combinar formas interactivas de aprendizaje colaborativo usando las nuevas tecnologías. Si bien, para asegurar la eficacia del aprendizaje, el docente tiene que (Bernard y Bachu, 2015):

- Conocer los conceptos previos de sus alumnos relacionados con la materia objeto de aprendizaje.
- Formular el problema, para orientar al alumno en la estructuración mental del mismo.
- Diseñar la solución: organización y refinamiento siguiendo un módulo lógico de especificación. Planificación de las soluciones (descubrimiento de estrategias, descomposición en metas).
- Modelar los datos.
- Realizar una diagnosis del error.
- Realizar la evaluación (análisis crítico, evaluación y revisión).
- Dar un *feedback* sobre las respuestas de aprendizaje.

El CSCL es uno de los elementos de innovación implicados en la mejora del proceso de enseñanza-aprendizaje fundamentado en el uso de las nuevas tecnologías (Biggs, 1985; Bernard y Bachu, 2015; García-Peñalvo, Alarcón, y Domínguez, 2019). Las estrategias que van a permitir una buena resolución de los problemas se pueden sintetizar en los siguientes componentes relacionados con el diseño de las materias (Mayer, 1998; 2003; de Raadt, Watson, y Toleman, 2009):

1. Comprensión correcta del problema que permita responder a las cuestiones qué tengo que hacer y para qué voy a hacerlo.

2. Establecimiento de los pasos en la resolución de problemas y la correcta secuenciación de estos (planificación).
3. Análisis de si existe más de un posible algoritmo de resolución y elegir el mejor proceso de resolución justificando la elección.
4. Identificación de las estrategias implicadas en cada uno de los pasos del proceso de resolución.
5. Verificación de la solución aportada.

Las plataformas colaborativas (*Learning Management System –LMS–*), si contienen un buen diseño pedagógico, facilitan el desarrollo de aprendizajes eficaces, ya que incrementan la co-construcción del conocimiento. Asimismo, estos entornos mejoran las estrategias de autorregulación e incrementan el uso de las estrategias metacognitivas, ya que aumentan la motivación del alumno hacia el objeto de aprendizaje. También facilitan la construcción del aprendizaje desde el uso de la monitorización y la corrección del error. Además, permiten dar un *feedback* sobre el desarrollo del conocimiento y sobre el uso de las estrategias para la resolución de tareas.

En síntesis, los CSCL incrementan la autorregulación, la planificación, el uso de las estrategias y la motivación hacia el aprendizaje, ya que facilitan la implementación de las rutinas de resolución y ponen en marcha habilidades cognitivas relacionadas con las estrategias cognitivas básicas (*basic cognitive skills*) y con las habilidades metacognitivas. Para Mayer (1998) ambas habilidades son necesarias para establecer una rutina de resolución de problemas y una reflexión sobre las estrategias de resolución.

1.3.2. Computer Based Learning Environment (CBLE)

El CBLE proporciona a los estudiantes la posibilidad de desarrollar un andamiaje para el aprendizaje en procesos clave, lo que potencia el aprendizaje eficaz. La utilización de los recursos hipermedia permite el aprendizaje auto-regulado (SRL), ya que estos estructuran la planificación, la monitorización, el control y la reflexión. Todo lo cual, incrementa la participación del aprendiz en el proceso de aprendizaje. Los medios tecnológicos facilitan, pues, la adaptación a las características individuales de aprendizaje de los sujetos. Asimismo, inducen un *feedback* directo e inmediato sobre la resolución de una tarea, tanto del proceso, como en el producto. Investigaciones en este campo indican que el aprendizaje utilizando

las nuevas tecnologías facilita la comprensión y la construcción del conocimiento, siempre que produzca un cambio en la estructura cognitiva y un incremento en el conocimiento declarativo y procedimental a través del uso de las estrategias de autorregulación (Cerezo, Sánchez-Santillan, Paule-Ruiz, y Nuñez, 2016). Las investigaciones indican que los sujetos que aprenden con un andamiaje de dificultad creciente en las tareas desarrollan un sistema cognitivo más complejo y autónomo en los procesos de resolución de las tareas (Azevedo, 2005). La estructuración escalonada del aprendizaje, especialmente en entornos de resolución de problemas, facilita al aprendiz la secuenciación tanto de las metas como de los pasos necesarios para la resolución de las tareas (Höök y Eckerdal, 2015; Winne, 2014), el uso del SRL y el incremento de la motivación (Segedy y Biswas, 2015). Una buena estructuración de CBLE tiene que contemplar:

1. Planificación, que incluye el establecimiento de metas, la activación del conocimiento previo y el reciclado de las metas en la memoria de trabajo o usando en los tres casos el condicional «establecer las metas, activar el conocimiento previo y reciclar las metas en la memoria de trabajo».
2. Actividades de monitorización, utilizando actividades dinámicas hipermedia que incluyen:
 - Sentimientos sobre el conocimiento (*Feeling of knowing*).
 - Juicios sobre el aprendizaje (*Judgment of Learning*).
 - Seguimiento del progreso (*Monitoring progress*) a través de la comprobación de las metas, contiene evaluación, identificación adecuada de la información, evaluación del contenido de las respuestas a las metas y las auto-preguntas.
 - Discriminación de estrategias de aprendizaje efectivas y no efectivas. Estrategias de aprendizaje que incluyen: hipótesis, coordinar las distintas fuentes de información, inferencias, reglas nemotécnicas, diagramas, resúmenes, estrategias de búsqueda, seleccionar nuevas fuentes de información, elaboración del conocimiento entre otras.
3. Procedimientos de detección de las dificultades de la tarea: propuesta de actividades con distinto grado de dificultad.
 - Desarrollo de conductas de búsqueda de ayuda (*help-seeking behaviour*).
 - Tiempo y esfuerzo en la planificación.
 - Análisis de la dificultad de la tarea.
4. Interés hacia la tarea o el contexto de dominio de la tarea.

1.3.3. Diferencias y similitudes entre CBLE y CSCL

En síntesis ambos entornos (CSCL y CBLE) son espacios virtuales interactivos de aprendizaje que tienen en común el partir de un estudio de los conocimientos previos de los estudiantes y proponer un diseño que se ajuste a las necesidades particulares de cada uno de ellos. El CSCL se centra en la utilización del aprendizaje autorregulado y el feedback proponiendo rutinas de resolución desde una monitorización del proceso de aprendizaje. El CBLE contempla también el aprendizaje autorregulado y monitorización, si bien utiliza más recursos hipermedia para generar un andamiaje en orden progresivo de dificultad para la construcción de los conocimientos por parte del aprendiz. Ambos entornos potencian la puesta en marcha de las habilidades cognitivas y metacognitivas en el aprendizaje.

1.4. Metacognición e Inteligencia Artificial

En los últimos años la metacognición se está utilizando como referente para el diseño de los procesos de Inteligencia Artificial con el fin de incrementar el nivel de autonomía de los sistemas inteligentes (Li, Wang, y Bai, 2015). Si bien, diseñar sistemas con habilidades metacognitivas es una difícil tarea, debido al número y complejidad de los procesos que este sistema contempla en el ser humano. Un dominio específico de utilización es el lenguaje visual M++. En M++ la especificación del nivel metacognitivo (*object-level*) y el metacognitivo (*meta-level*) son la base de a un metamodelo que se ha configurado acorde a ambos conceptos. M++ permite la generación de diagramas metacognitivos en un editor visual que se denomina *MetaThink*. Para la validación del sistema se utilizan dos técnicas: un estudio empírico y un modelo de rastreo. Se ha demostrado que la anotación M++ es útil en el proceso de modelización de componentes en sistemas inteligentes. La técnica de rastreo consiste en MOF (Modelo de Organización de funciones) basado en el metamodelado (Caro, Josyula, Jiménez, Kennedy, y Cox, 2015). No existe un modelo único de diseño computacional de los procesos metacognitivos, actualmente se ha puesto el énfasis en el DSVL (*Domain-Specific Visual Language*) para los procesos metacognitivos, éste incluye una sintaxis abstracta y también compleja.

MOF-based *metamodel* contiene componentes de metacognición y se compone de tres tipos de elementos:

- Elementos estructurales. Contiene elementos funcionales y elementos básicos. Se compone de dos niveles: cognitivo (razonamiento sobre la resolución de problemas) y metacognitivo: *object-level* y *meta-level* (nivel de representación de del sistema de razonamiento, monitorización y control del propio aprendizaje y de los procesos de razonamiento).
- Elementos funcionales. Son tareas que permiten el razonamiento y la toma de decisiones
- Elementos básicos, elementos que participan e interaccionan en el meta model metacognitivo.

Los nexos de relación son de:

- Asociación ()
- Generalización ()
- Agregación ()

El punto de arranque (*main*) de los elementos funcionales en el metamodelo son: tareas de razonamiento (TR) y de meta-razonamiento (MT). Las primeras son las acciones que permiten el procesamiento (transformación, reducción, elaboración, almacenamiento y recuperación) y su aplicación al conocimiento y la toma de decisiones para cumplir los objetivos del sistema y las segundas permiten explicar los errores en algunas de las tareas de razonamiento y pueden usar la selección entre algoritmos metacognitivos que conforman el razonamiento del *object-level*.

En resumen, el objeto del nivel de especificación (*object-level specification*), se compone de tareas de razonamiento, de las huellas del razonamiento y de las clases de las huellas de razonamiento. Los datos computacionales contienen una salida que genera la tarea como datos parciales desde el procesamiento computacional y de las tareas de razonamiento disponibles desde un juego de estrategias que están asociadas a las metas. La revisión continua de las estrategias de razonamiento y de las estrategias empleadas en cada paso son actualizadas de forma constante en el proceso de toma de decisiones y pertenecen al *meta-level*. El *meta-level* de especificación, describe los mecanismos de meta-razonamiento, implican la monitorización introspectiva, detectan los fallos en el razonamiento a nivel de *object-level* y comportan mucha información de retroalimentación que va a permitir tomar decisiones efectivas a nivel de *meta-level control*, y el *meta-level control* (Cuba-Ricardo, Serrano-Rodríguez, Leyva-Figueroa, y Mendoza-Tauler, 2015) (ver Figura 6).

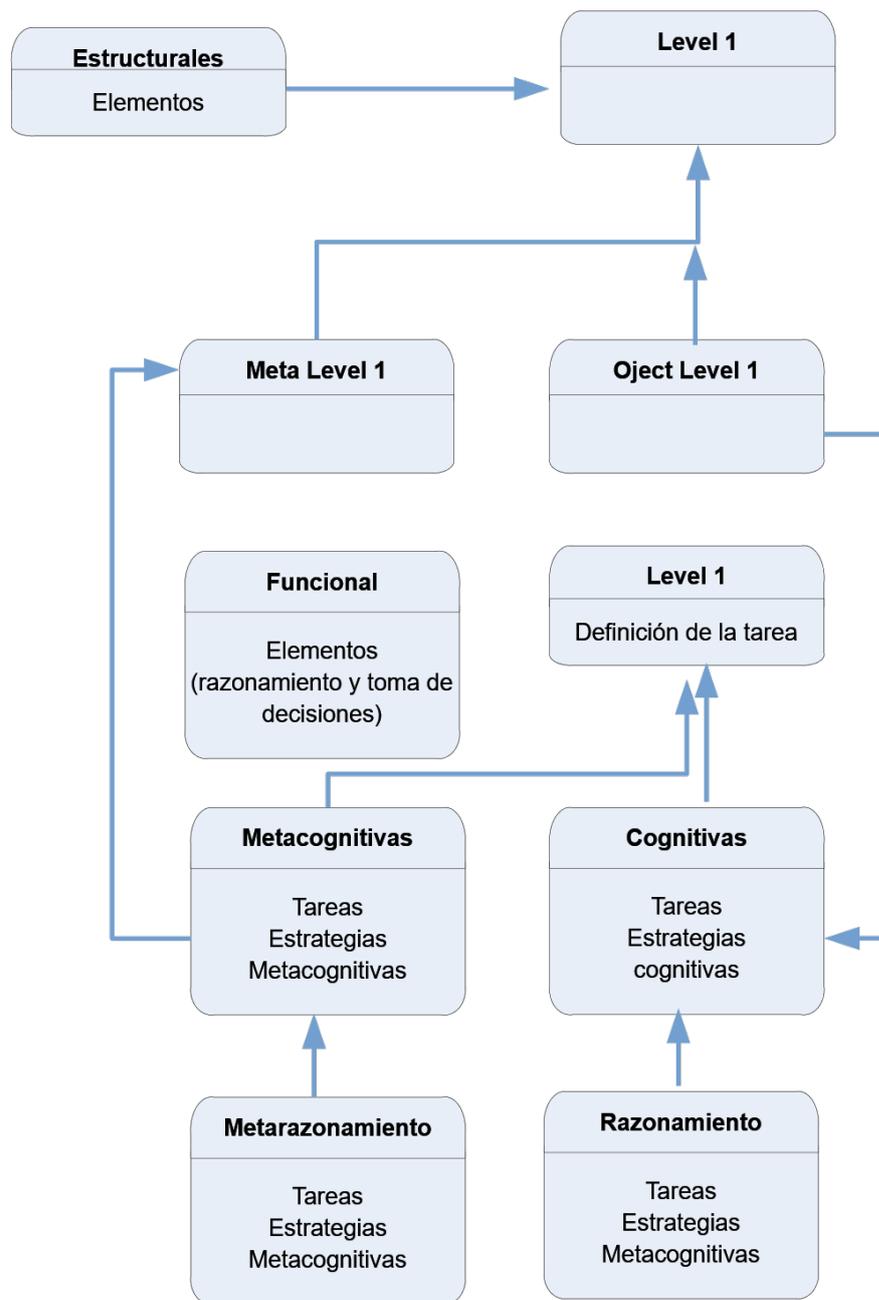


Figura 6

Relaciones entre los elementos funcionales y los estructurales adaptado de Cox y Raja (2012)

1.5. Aprendizaje Basado en Proyectos: desarrollo en entornos virtuales

El Aprendizaje Basado en Proyectos (ABP), es un método de enseñanza-aprendizaje que se basa en el constructivismo (Hmelo-Silver, 2004). El primer autor en proponerlo fue

John Dewey a finales de la década de 1890 (Douglas y Stack, 2010). La filosofía de Dewey se centraba en el aprendizaje desde situaciones de la vida real. Sus ideas fueron desarrolladas por Kilpatrick a principios de 1900 en su libro «El Proyecto Método» (1918). Desde entonces, el ABP ha sido elaborado en detalle y aplicado a distintas materias en diferentes situaciones (Habók y Nagy, 2016 p. 5). El ABP se puede definir como la instrucción centrada en el estudiante que se produce durante un período de tiempo prolongado durante el cual los estudiantes seleccionan, planifican, investigan y realizan un producto, presentación o desarrollo, que responde a una pregunta o a un problema de investigación (Holm, 2011 p.1). Otra definición de ABP es, un método sistemático de enseñanza que ayuda a los estudiantes a engranar sus conocimientos y habilidades a través de tareas y de preguntas que están cuidadosamente diseñadas (Markham, Larmer, y Ravitz, 2003, p. 4). El ABP se centra en el proceso de enseñanza-aprendizaje, en la interacción del docente y los alumnos, en la construcción de aprendizajes profundos desde el diseño de tareas basadas en la investigación orientadas desde la práctica del conocimiento de la disciplina y dirigidas a través de preguntas de reflexión. Todo ello desde la realización de un *feedback* formativo continuo de los procesos y de los productos de aprendizaje en un marco de aprendizaje colaborativo (Doise, Mugny, y Perret-Clermont, 1975) (ver Figura 7).

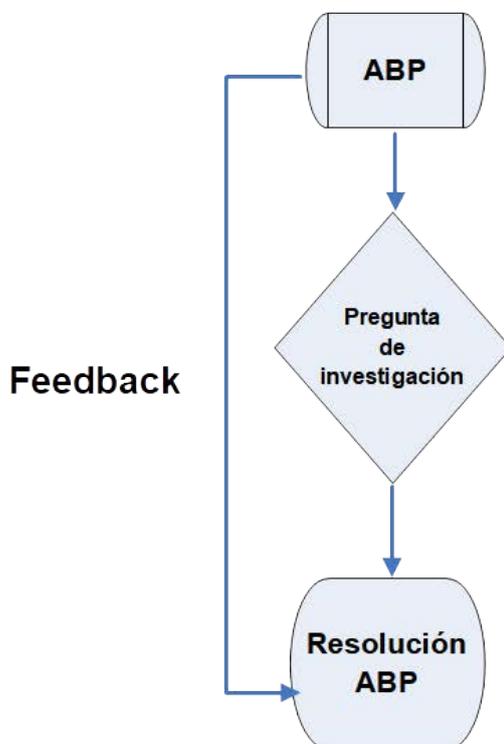


Figura 7

Proceso de ejecución del ABP

Investigaciones más recientes (Hovey y Ferguson, 2014) señalan que se pueden diferenciar diferentes tipos de ABP:

- Aprendizaje basado en la investigación.
- Aprendizaje basado en un problema.
- Aprendizaje basado en la actividad.
- Aprendizaje por descubrimiento.

Asimismo, estudios actuales (Hmelo-Silver y DeSimone, 2013; Lajoie et al., 2015) señalan que un factor importante para el éxito en la elaboración de los ABP es el uso de plataformas interactivas como el CSCL. Dichos entornos son instrumentos que facilitan la co-regulación y el desarrollo de las estrategias metacognitivas en la resolución de tareas. Se ha comprobado que los estudiantes emplean más estrategias metacognitivas de aprendizaje si utilizan los entornos virtuales colaborativos (Bannert, Sonnenberg, Mengelkamp, y Pieger, 2015). Así pues, los CBLE permiten el diseño y la aplicación de programas estructurados de desarrollo y activación de las estrategias metacognitivas y de las estrategias de apoyo y transferencia del aprendizaje. Lo que facilita el desarrollo de la autorregulación del aprendizaje (*Self-Regulated Learning* -SRL-) (Bannert, Reimann, y Sonnenberg, 2014), ya que estructuran de forma escalonada las submetas necesarias, para lograr un aprendizaje profundo y exitoso desde aproximaciones sucesivas a la meta (Azevedo, Johnson, Chauncey, y Graesser, 2011).

Una de las posibles razones es que la guía en los procesos de navegación a través de las plataformas interactivas facilita el éxito en el aprendizaje. El soporte diseñado desde planteamientos metacognitivos potencia el incremento de la competencia de aprendizaje del alumno, ya que aumentan el uso de las estrategias de planificación (Sáiz y Montero 2016; Marinho-Araújo y Almeida, 2016). Un resumen del procedimiento se expone en la Tabla 2.

Tabla 2

Relación entre los conceptos de conocimiento de la cognición, regulación de la cognición y conocimiento metacognitivo en las estrategias de orientación

Estrategias del docente	Estrategias del alumno
Estrategias de explicación	Análisis de la información (consulta de la misma en la plataforma).
Estrategias de control de la adquisición de la explicación (análisis de los fallos en comprensión y de los conocimientos previos necesarios para la comprensión del tema).	Reflexión sobre los conocimientos previos que exige la materia y determinar los que se poseen y los que no. Análisis de los conceptos que se han asimilado y los que no.
Diseño de prácticas que apoyen la comprensión de los conocimientos teóricos.	Ejecución de la práctica. Explicitación sobre las dudas.
<i>Feedback</i> del docente sobre la ejecución de la práctica.	Análisis del <i>feedback</i>
Trabajo desde ABP realización de un proyecto basado en la aplicación de los conocimientos teóricos.	Ejecución del proyecto. Explicitación de las dudas.
<i>Feedback</i> continuo a lo largo de la realización del proyecto (establecimiento de entregas parciales, revisión y <i>feedback</i> sobre las mismas).	Análisis del <i>feedback</i> .

Lo que ha llevado a muchos investigadores a analizar el efecto del *e-learning* desde el uso de los recursos hipermedia en las respuestas de aprendizaje de los estudiantes (Norman y Furnes, 2016). Los resultados indican que la utilización de los medios hipermedia facilita el desarrollo de un aprendizaje más profundo y de mayor calidad (de Kock, 2016), ya que la planificación sistemática a través de programas de ordenador, que combinan tanto los elementos de audio como de vídeo, facilita el desarrollo de las estrategias metacognitivas en la resolución de problemas. Dichas plataformas pautan el objetivo, la planificación en el desarrollo de la resolución, la evaluación de los pasos de resolución y la evaluación del resultado final desde un respeto al ritmo de aprendizaje del alumno (Matwin y Mielniczuk, 2016).

1.6. El proceso de enseñanza-aprendizaje en LMS

Si bien, el mero uso de los LMS no garantiza mejores resultados de aprendizaje. Dicho uso está condicionado por un lado al diseño que el docente realiza de las actividades de aprendizaje y por otro al tipo de retroalimentación que dicho sistema proporciona sobre las evidencias de aprendizaje. Estudios recientes (Yamada y Hirakawa, 2015; Park y Il-Hyun, 2016) indican que para un desarrollo satisfactorio del proceso de enseñanza-

aprendizaje en los LMS se precisa una formación sobre su uso tanto en los docentes como en los estudiantes. Asimismo, dichos autores han encontrado diferencias significativas atendiendo a variables como las características de los docentes, de los estudiantes, de la estructuración y del diseño de las asignaturas.

Otro aspecto relevante en el proceso de aprendizaje en los LMS es la detección temprana del estudiante en riesgo de obtener resultados académicos pobres. Estudios recientes (Strang, 2016; Zacharis, 2015) analizan la relación entre el uso de los LMS y los patrones de aprendizaje de los estudiantes en la plataforma. Las técnicas de análisis de regresión permiten detectar los patrones conductuales exitosos y los de riesgo. Dichos patrones explican hasta el 52% de la varianza en los resultados de aprendizaje. Estos estudios están validados mediante el uso de Técnicas de Minería de Datos apoyadas en la utilización de herramientas como *Waikato Environment for Knowledge Analysis* (WEKA). Las conductas de aprendizaje que se han considerado referenciales en un aprendizaje exitoso son (Cerezo et al., 2016; Zacharis, 2015):

- El esfuerzo.
- El tiempo que utilizan en realizar las tareas.
- El tiempo que utilizan en el trabajo en los contenidos teóricos.
- Los resultados en los test de auto-evaluación (*quizzes efforts*).
- El tiempo empleado en las discusiones en los foros.
- La calidad de las discusiones en los foros (tipo de mensaje y longitud del mismo).
- Tiempo empleado en analizar el *feedback* realizado por el docente.
- El número y tipo de los mensajes enviados.
- La frecuencia de uso del LMS.
- La lectura de los mensajes en el foro.
- La contribución a la creación de contenidos.
- Los archivos vistos.
- El tiempo de entrega de las actividades.

1.6.1. Logs, Analíticas de aprendizaje (LA) y Minería de datos educativos (EMD)

Cuando los distintos participantes del proceso de aprendizaje interactúan a través de un LMS se generan una serie de *logs* o archivos de registro que capturan cada una de las interacciones. Dichos *logs* son susceptibles de ser analizados. El uso de Técnicas de Minería de Datos va a permitir aislar patrones o extraer nueva información desde el análisis de grandes conjuntos de datos. Cuando estas técnicas se emplean con datos relativos al aprendizaje se habla de Analíticas de Aprendizaje (*Learning Analytics* -LA-) o Minería de Datos Educativos (*Educational Data Mining* -EDM-). Estos conceptos están muy relacionados entre sí, el primero se centra más en el entendimiento del proceso de aprendizaje, y el segundo en los modelos que permiten analizar los datos (Baker, 2014; Baker y Inventado, 2014). Dentro de los LMS, uno de los más utilizados es Moodle. Se trata de un LMS de estructura modular. Permite utilizar distintos recursos para distintos perfiles de alumnos (individual y grupo) y profesores. También, posibilita realizar diferentes actividades y acciones de aprendizaje (foros de discusión, cuestionarios, talleres, wikis, acceso a repositorios, etc.) y utilizar metodologías docentes innovadoras como el Aprendizaje Basado en Proyectos (ABP). Las conductas interactivas que se pueden analizar sobre este tipo de LMS, son (Yücel y Usluel, 2016).

- Interacción estudiante-estudiante.
- Interacción estudiante-docente.
- Interacción estudiante-contenido.
- Interacción estudiante-sistema.
- Interacción docente-sistema.

Yücel y Usluel (2016) señalan que es importante considerar el tipo, la cantidad y la calidad de la interacción. Cada una de estas conductas interactivas deja reflejado su uso en los ficheros de registro o *logs*. Para facilitar su análisis, Moodle permite la extracción de estos archivos que es donde han quedado almacenadas las interacciones de los miembros de la comunidad de aprendizaje en los distintos eventos. El análisis de dichos archivos va a proporcionar mucha información acerca de las conductas de aprendizaje, estudiantes, y de docencia, profesor. Si bien, la información que se puede obtener de los archivos de registro de Moodle es muy extensa, es por lo que el empleo de las Técnicas de Minería de Datos va a permitir la extracción de la información que se precise en cada caso (Iglesias-Pradas, Ruiz-

de-Azcarate, y Agudo-Peregrina, 2015). Así pues, la EDM desarrolla técnicas y modelos que van a facilitar el conocimiento de los patrones de conducta de aprendizaje de los alumnos y de las interacciones entre ellos (docente-estudiante, estudiantes-estudiantes) (Gibson y de Freitas, 2016). Asimismo, proporcionan métodos de extracción de la información en tiempo real. Todo lo cual apoya a los procesos de evaluación continua (formativa). La EDM puede dirigirse a distintos perfiles (Romero y Ventura, 2007; Romero, Espejo, Zafra, Romero, y Ventura, 2013):

- Orientados hacia los estudiantes (*oriented towards students*). Está enfocada a las tareas de aprendizaje y el objetivo es el de mejorar el proceso de aprendizaje en los estudiantes.
- Orientados hacia el docente (*oriented towards educators*). El objetivo es el de proporcionar retroalimentación para la instrucción, evaluar la estructura y los contenidos del curso, analizar elementos que han sido efectivos en los procesos de aprendizaje, clasificar el tipo de alumnos y ver las necesidades de guía y de monitorización del aprendizaje. Todo ello, con el fin de conocer los patrones más habituales en su aprendizaje así como la frecuencia de errores para determinar las actividades más efectivas.
- Orientados hacia los responsables educativos (*oriented towards academics responsible and administration*). El objetivo es el de dar información a la institución que ayude a mejorar las plataformas de aprendizaje. Como se ha indicado anteriormente las LA se relacionan con las técnicas de EDM. Si bien, las primeras están enfocadas al estudio y análisis del proceso de aprendizaje (Agudo-Peregrina et al., 2014; Chatti, Dyckhoff, Schroeder, y Thüs, 2012) en concreto, las técnicas de LA investigan, entre otras, las respuestas a las siguientes cuestiones:
 1. Qué datos analizar (*what*). La información que se registra en los LMS es abrumadora por ello se precisa de la detección de patrones de análisis de datos.
 2. Para quién se proporciona esa información (*who*). Es importante discriminar el colectivo al que va a ir dirigido el análisis (estudiantes, profesores, tutores/mentores, administradores educativos, etc.). Los estudiantes estarán interesados en espacios de aprendizaje más efectivos. Los profesores en cómo hacer más eficaces sus prácticas de enseñanza y en ofrecer el apoyo que sus estudiantes necesitan, y las instituciones en detectar a los alumnos en riesgo con

el fin de incrementar el éxito de sus estudiantes. Por todo ello, cada vez más, se plantea el implementar herramientas dentro de los LMS que ofrezcan el análisis de los datos para que los no expertos puedan estudiar las distintas situaciones. Estas técnicas ofrecen un *feedback* orientado hacia la tarea (*goal oriented feedback*) que permite al usuario de la plataforma la reflexión sobre sus acciones y orientan la toma de decisiones.

3. Por qué se proporciona la información (*why*). Hay diferentes objetivos en función del rol de usuario. Las LA incluyen el análisis de monitorización (*monitoring analysis*), es decir, el seguimiento de los estudiantes con el fin de generar informes para el profesor y/o para la institución. Esta información va a facilitar la evaluación del proceso de aprendizaje por parte del docente con el fin de mejorar el entorno de aprendizaje y ofrecer ayudas al estudiante que le permitan incrementar los resultados). Así como la predicción del conocimiento del estudiante y de los resultados de aprendizaje para detectar al alumno en riesgo y proporcionarle, en su caso, las ayudas necesarias para que logre un aprendizaje eficaz. Además de implementar la tutorización (proceso de orientación del aprendizaje) y la mentorización (plan concreto de ayuda personalizada en cuestiones de planificación y supervisión y preparación de nuevos retos de forma específica para cada alumno en función de sus necesidades); la evaluación y el *feedback* (facilita procesos de auto-evaluación que permitan al estudiante su éxito en el aprendizaje). Se realizan informes de *feedback* inteligente que se proporcionan tanto al docente como al estudiante).
4. Cómo (*how*) se proporciona la información. Los métodos para la detección de los patrones ocultos (*hidden patterns*) de aprendizaje en los LMS se pueden realizar a través de métodos estadísticos, métodos de visualización de datos y Técnicas de Minería de Datos.
 - *Métodos estadísticos*: los LMS permiten la extracción de informes basados en la interacción del profesor y de los alumnos en la plataforma que analizan el tiempo en la plataforma (*time online*), número de visitas (*total number of visits*), número de visitas por elemento (*number of visits per page*), distribución del tiempo de visitas (*distribution visits over time*), frecuencia de réplicas, etc. Los análisis estadísticos que proporcionan son medias (*M*) y desviaciones típicas (*DT*).

- *Información de visualización*: establecen informes fáciles de interpretar por los usuarios de las distribuciones de los datos. Entre las distintas técnicas se pueden diferenciar, entre otros, los gráficos de área de barras, histogramas, mapas de calor, diagramas de dispersión (en 2D y en 3D).
- *Minería de datos*: se generan en categorías de aprendizaje supervisado (clasificación y regresión), no supervisado (agrupamiento o *clustering*) y reglas de asociación de datos (Arnaiz-González, Díez-Pastor, Rodríguez, y García-Osorio, 2016a).
 - Aprendizaje supervisado

Dentro de él se pueden diferenciar técnicas de clasificación y de regresión y otras que se utilizan indistintamente como son las técnicas de Métodos multi-modelo (*Ensemble Methods*), Árboles de decisión (*Decision Trees*) y Redes Neuronales (*Neural Networks*) (Broomhead y Lowe, 1988).

Clasificación, serían procesos de hallazgo de una función o de un modelo que distingue las clases de datos. El objetivo es el de predecir las clases de objetos cuya etiqueta de clase se desconoce. Incluye los árboles de decisión, redes neuronales, clasificación bayesiana y máquinas de soporte vectorial [*support vector machines* (SVM)] junto con los modelos de predicción de vecinos más cercanos [*k-nearest neighbor classification* (*kNN*)]. La clasificación predice niveles de categoría.

Regresión, esta técnica se utiliza para realizar las funciones de predicción. Entre ellas, se pueden diferenciar *Linear Regression/Support Vector Machine* para regresión (SVR).

La diferencia fundamental entre las técnicas de clasificación y las técnicas de regresión es que en las primeras el número de valores a predecir es finito y no existe relación de orden entre las clases y en las de regresión el valor de predicción puede tomar valores infinitos y sí existe una relación de orden.

Ensemble Methods, son meta-algoritmos que combinan varias técnicas de aprendizaje automático en un modelo predictivo con el fin de disminuir la varianza, el sesgo o mejorar las predicciones.

- Aprendizaje no supervisado

Dentro de él se incluyen las técnicas de *clustering* o agrupamiento. En contraste con la clasificación en esta técnica el nivel de clase en cada objeto de entrenamiento o instrucción no es conocido. Asimismo, esta técnica organiza los datos en clases desde el criterio de inclusión de clúster semejante vs. clúster diferente. El criterio es cómo los datos se sitúan en el espacio, se basan en una función de distancia. Los métodos de agrupamiento se clasifican en: métodos de partición (*partitioning methods*), métodos de jerarquización (*hierarchical methods*) y métodos de densidad (*density-based methods*).

En los métodos de partición cada partición representa un clúster. La técnica comienza con una partición aleatoria seguidamente se aplica un proceso iterativo de relocalización que mejora la partición y mueve los objetos desde un grupo a otro. Estas técnicas utilizan métodos heurísticos populares como son el algoritmo de *k-mean* (cada clúster representa desde el valor medio de los objetos en el clúster) y el algoritmo de *k-medoids* (cada clúster es representado desde uno de los objetos del clúster).

En los métodos de jerárquización se crea una jerarquía de descomposición (árbol de clúster). El uso de cada división aglomerativa o divisiva depende de si la jerarquía de descomposición se realiza de forma ascendente (*bottom-up* o *fusión-merging*) o de forma descendente (*top-down* o *splitting*).

Finalmente, en los métodos de densidad se basan en la detección de las áreas en las que existen concentraciones de puntos y aquellas en las que la concentración es muy escasa. Dentro de estos algoritmos se pueden diferenciar DBSCAN (*Density Based Spatial Clustering of Applications with Noise*) (Ester, Kriegel, Sander, y Xu, 1996), OPTICS (*Ordering Points To Identify the Clustering Structure*) (Ankerst, Breunig, Kriegel, & Sander, 1999), DENCLUE (DENsity-based CLUstEring) (Campello, Moulavi y Sander, 2013), CLIQUE (Clustering in QUEst) (de Medeiros et al., 2008) y SNN (Shared Nearest Neighbor) *density-based clustering* (Liu, Wang, y Yu, 2018).

En resumen, la técnica de EDM es multidisciplinar, en ella convergen técnicas de construcción de algoritmos, redes neuronales artificiales (*Artificial Neural Networks*), aprendizaje basado en instancias (*instance-based learning*), aprendizaje bayesiano, etc. Dichas técnicas pueden utilizar diferentes procedimientos de análisis (Arnaiz-González, Díez-Pastor, Rodríguez, y García-Osorio, 2016a). Dichos procedimientos se pueden agrupar en técnicas de análisis de clústeres (*clustering*), técnicas de clasificación y detección de elementos extremos (*outlier detection*), minería de reglas de asociación y minería de patrones de secuencias y texto (*association rule mining and sequential pattern mining and text mining*) (Romero y Ventura, 2007). Por lo que el uso de las diferentes técnicas en EDM dependerá de los objetivos del análisis de las tareas. No obstante, los investigadores en este campo buscan la predicción de los resultados para proporcionar las recomendaciones particulares en cada caso. El reto de LA se centra en el análisis de tareas que permita proporcionar a los docentes de información relevante para poder intervenir de forma individualizada en el proceso de aprendizaje de manera inmediata y efectiva. Asimismo, otro aspecto que hay que considerar es que los patrones de comportamiento en la plataforma dependen del tipo de *Blended Learning (B-Learning)* utilizado por el docente (Margulieux, McCracken, y Catrambone, 2016). Por ejemplo, en *Replacement Blended (RB)*, se efectúa el *feedback* sobre las producciones de aprendizaje en la plataforma, en este entorno es esencial la participación en los foros de discusión. Sin embargo, en *Supplemental Blended (SB)*, se efectúa el *feedback* sobre las producciones *Face to Face (F2F)*, en este espacio no es importante el uso de los foros ya que la interacción se puede realizar de forma presencial (F2F) (Cerezo et al., 2016). Otra variable adicional es que no todos los estudiantes tienen el mismo proceso de aprendizaje en los LMS. Asimismo, otro indicador relevante es que la docencia en los LMS se puede diseñar de forma tradicional o de forma innovadora [*team based projects, online discussion forums y online Quizzes* (Park y Il-Hyun, 2016)]. El diseño de la docencia es esencial en los resultados de aprendizaje y en las conductas de los estudiantes en la plataforma (Sáiz, Marticorena, García-Osorio, y Díez-Pastor, 2017). Para completar los resultados obtenidos en el proceso de enseñanza-aprendizaje, también es necesario realizar una evaluación de la satisfacción de los usuarios (estudiantes y profesores) con el empleo de los LMS (Hornbæk, 2006). Los modelos de e-evaluación sugieren que hay distintas variables que influyen en su uso como son las estrategias de aprendizaje empleadas por los estudiantes (Pitigala-liyanage, Lasith-Gunawardena, y Hirakawa, 2016; Sáiz, Montero, Bol, y Carbonero, 2012), el medio en el que se desarrolla el aprendizaje (Harrati,

Bouchrika, Tari, y Ladjailia, 2016), el diseño que el docente realiza de la enseñanza (Cerezo, et al., 2016; Sáiz et al., 2017), y el uso de metodologías docentes activas (Sáiz y Bol, 2014; Sáiz y Montero, 2016). También, se ha encontrado distintos patrones y resultados de aprendizaje dependiendo del tipo de e-evaluación efectuada (Sáiz y Arnaiz, 2017).

En síntesis, el uso de las metodologías anteriormente descritas va a permitir detectar los patrones de comportamiento en la plataforma de los estudiantes y del docente desde el estudio de los *logs*. Asimismo, las EMD facilitarán el estudio del comportamiento y el desarrollo de Ciencia Cognitiva (Jones, 2017).

1.7. Personalización del aprendizaje y e-orientación

Estudios recientes en Educación Superior indican que las metodologías docentes deben dirigirse a formas más personalizadas de interacción con el estudiante (Gairín, Feixas, Guillamón, y Quinque, 2004; Sáiz y Montero, 2016). Por ello, la enseñanza tiene que ser personalizada para poder ajustarse al ritmo de aprendizaje de los estudiantes (Sáiz y Román, 2011). El docente, pues, tiene que realizar un proceso de reflexión sobre su propia práctica (Gimeno Sacristán, 2008). Es por lo que la Educación Superior se encuentra en un momento de cambio derivado de las nuevas demandas de la sociedad del conocimiento. Por ello, es necesario que el profesor universitario realice una tarea de orientación en el proceso de aprendizaje de cada alumno. En España esta función se ha normativizado en el Estatuto del Estudiante Universitario (Real Decreto 1791/2010, de 30 de diciembre, BOE nº 318). Dicho Estatuto en el artículo 19 indica que los estudiantes recibirán orientación y seguimiento de carácter transversal sobre su titulación, en el artículo 20 refleja que la acción tutorial facilitará el proceso de: (a) transición y adaptación del estudiante al entorno universitario; (b) información, orientación y recursos de aprendizaje, (c) la configuración del itinerario curricular atendiendo también a las especificidades del alumnado con necesidades educativas especiales y la transición al mundo laboral, el desarrollo inicial de la carrera profesional y el acceso a la formación continua. También en el artículo 25.2 subraya la importancia de la evaluación continua del proceso de enseñanza-aprendizaje. Es por lo que la función de orientación puede ser entendida como un proceso de ayuda continuo y sistemático dirigido desde principios de prevención del abandono de la titulación por parte del estudiante (Bisquerra-Alzina, 1996). Así pues, la orientación al alumnado en la Universidad tiene que estructurarse desde el diseño de programas de enseñanza (Zabalza, 2003; Payo, et al., 2013) que potencien aprendizajes exitosos (Carbonero, Román, y Ferrer,

2013; Knight, 2005; Pozo y Del Puy, 2009; Sáiz y Román, 2011). Todo ello, reafirma a la función orientadora como valor inherente de la función docente (Alonso-Tapia, 1995; Zabalza y Cid, 2006) y dentro de ella la tutoría programada se ha mostrado efectiva para el seguimiento de la trayectoria de aprendizaje del estudiante (Sáiz y Román 2011). Muchas son las universidades que han desarrollado este tipo de programas (Román 2004a; Román 2004b), entre ellas destacan la Universidad de Washington, Universidad de Wisconsin-Madison, Oxford, Cambridge y dentro de nuestro país la Universidad de Alcalá de Henares, Alicante, Autónoma de Madrid, Burgos, Oviedo, Granada, Murcia, Navarra, Rovira i Virgili, Zaragoza, Valencia. Recientes estudios (Sáiz, Marticorena, y García Osorio, 2017) parecen confirmar que la orientación personalizada al estudiante universitario facilita el aprendizaje eficaz. En este campo las nuevas tecnologías han abierto un nuevo entorno para el aprendizaje (Gros y García Peñalvo, 2016). Los LMS como Moodle, y en concreto UBUVirtual¹, permiten respetar el ritmo de aprendizaje del alumno y realizar un seguimiento individualizado del proceso de enseñanza-aprendizaje. La extracción y análisis de los registros o *logs* facilita la detección del alumno en riesgo. Recientes estudios (Sáiz et al., 2017) indican que el seguimiento personalizado del alumno incrementa sus resultados de aprendizaje y su frecuencia de interacciones en la plataforma. Asimismo, dicho seguimiento predice los resultados de aprendizaje globales de los estudiantes en un 61.3% y los sus patrones de comportamiento en un 56.1%.

1.7.1. Feedback y tutoría programada en la Universidad

Recientes investigaciones (Wentzel y Watkins, 2011; Sáiz y Payo, 2012) señalan que los contextos de aprendizaje estructurados hacia la tutoría programada incrementan la motivación del estudiante hacia la materia de estudio. La tutoría programada se ha relacionado con el desarrollo de las competencias cognitivas y metacognitivas y dentro de ellas con las habilidades de resolución de problemas y de razonamiento. La tutorización, puede facilitar el desarrollo de una orientación personal y académica siempre que se estructure hacia la generación de auto-preguntas, la clarificación de aspectos prioritarios (apoyo en la toma de decisiones) y la realización de un *feedback* sobre el proceso de enseñanza-aprendizaje (Sáiz y Montero, 2015). Todo ello va a facilitar el desarrollo de las

¹ Es un Learning Management System basado en Moodle que se utiliza en la Universidad de Burgos tanto en docencia presencial como no presencial.

competencias (conceptuales y procedimentales) que el sistema universitario y posteriormente el entorno laboral exigen al estudiante y futuro egresado (Sáiz y Román, 2011). Los beneficios de este tipo de programas se traducen en la calidad del aprendizaje (Biggs, 2005). Desde este planteamiento la función del docente tiene que ser personalizada y preventiva para evitar el fracaso y el posible abandono del estudiante.

1.7.2. Detección del alumno en riesgo de fracaso a través de las LMS

La detección del alumno en riesgo académico tiene que ser un objetivo prioritario para el profesor y para las instituciones universitarias (Sáiz et al., 2017). Para poder efectuar una detección eficaz del alumno en riesgo se ha mostrado efectiva la ayuda de los *Learning Management System* (LMS). Dichos sistemas van a permitir un registro y seguimiento de los patrones conductuales de los estudiantes (Strang, 2016; Zacharis, 2015). Muchas de estas plataformas utilizan técnicas de LA o EDM, como las técnicas de análisis de regresión que ayudan en la detección y posterior predicción de las conductas exitosas y las de riesgo. Dichas conductas explican hasta el 52% de la varianza de los resultados de aprendizaje. Estos estudios están orientados desde el uso de Técnicas de Minería de Datos (Cerezo et al., 2016 y Zacharis, 2015).

En síntesis, la frecuencia y la sistematicidad en la interacción en la plataforma por parte del alumno junto con la realización de las actividades de auto-evaluación y las consultas medias por día son aspectos directamente relacionados con el aprendizaje eficaz (Yücel y Usluel, 2016).

En esta línea, recientes investigaciones (Sáiz, Marticorena, García-Osorio, y Díez-Pastor, 2019) revelan diferencias en la predicción de los resultados de aprendizaje atendiendo a la variable metodología docente (estructuración de la docencia y procedimientos de evaluación). El tipo de actividades y las pruebas de evaluación (preguntas tipo test, proyectos, presentaciones...) propuestas harán que el estudiante tenga distintos patrones de conducta en la plataforma.

El análisis de los *logs* o archivos de registro (capturas cada una de las interacciones) a través de Técnicas de Minería de Datos va a permitir al docente analizar los patrones de comportamiento de sus estudiantes y predecir al alumno en riesgo (Díez-Pastor, Arnaiz-González, Marticorena-Sánchez, y Sáiz-Manzanares, 2017). Asimismo, la intervención temprana en estos casos previsiblemente mejorará las respuestas de aprendizaje de los

estudiantes. También el uso de los LMS va a facilitar, especialmente en entornos universitarios, la estructuración de una enseñanza colaborativa lo que previsiblemente incrementará la motivación de los estudiantes hacia el proceso de aprendizaje (Bernard y Bachu, 2015; Järvelä et al., 2016; Malmberg et al., 2015; Sáiz y Marticorena, 2016). Recientes estudios confirman [Saiz et al (2017); Strang (2016); Zacharis (2015)] que el seguimiento de los patrones de aprendizaje de los alumnos en la plataforma facilita la discriminación del alumno en riesgo desde una varianza explicada del 67.2%.

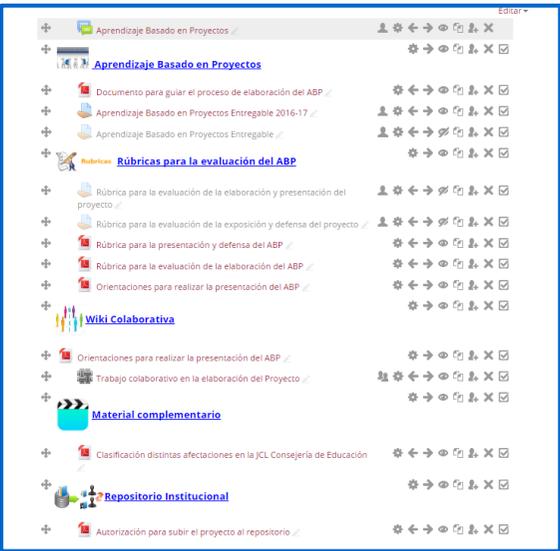
1.7.3. La personalización de aprendizaje

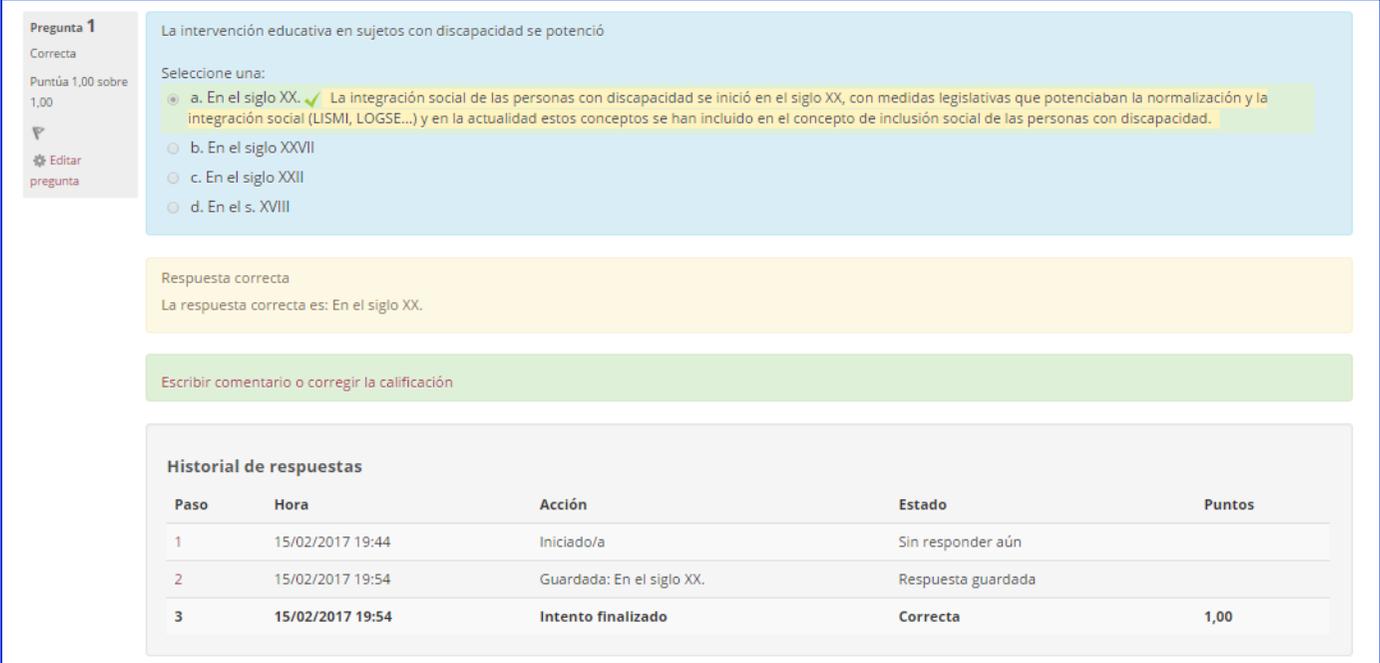
La personalización del aprendizaje es una de las áreas más importantes de investigación en computación dirigida al campo educativo (Narciss et al., 2014). Como ya se ha indicado anteriormente, con el fin de facilitar aprendizajes eficaces es importante monitorizar las acciones de los estudiantes y adaptarse sus características y necesidades con el fin de que estos desarrollen aprendizajes exitosos. Diferentes recursos tecnológicos han desarrollado y optimizado las intervenciones instruccionales del docente. Dentro de esas técnicas se encuentran los tutores cognitivos o tutores inteligentes que cuentan con rigurosas reglas basadas en modelos cognitivos de tareas paso a paso. Los recursos tecnológicos permiten el desarrollo de un aprendizaje por aproximaciones sucesivas a la meta, desde el uso de una retroalimentación sistemática en tiempo real sobre las producciones de cada estudiante lo que facilita un aprendizaje personalizado. Los sistemas de tutoría inteligente pueden registrar y monitorizar las interacciones con los estudiantes, qué es lo que conocen y cómo lo van conociendo. Asimismo, les pueden dar ayudas graduales en función de su ritmo de aprendizaje (Olney, Graesser, y Person, 2010). Un procedimiento para generar aprendizajes personalizados es la generación de diálogos que impliquen *feedback* y reflexión sobre las tareas, lo que lleva directamente a la construcción personalizada del conocimiento. Los recursos hipermedia combinan un *feedback* formativo y adaptativo con los sistemas de evaluación formativa. Estos sistemas posibilitan entrenar en el uso de qué estrategias seguir en el proceso de la resolución de una tarea o de un proyecto (*feedback-based learning technologies*). Asimismo, los entornos de aprendizaje virtual como Moodle, permiten el desarrollo de experiencias de aprendizaje en contextos de colaboración entre el estudiante, el profesor, los grupos y la tarea o los proyectos a elaborar (Ruggiero y Boehm, 2017). En concreto, en los estudios que fundamentan este proyecto de investigación se ha desarrollado

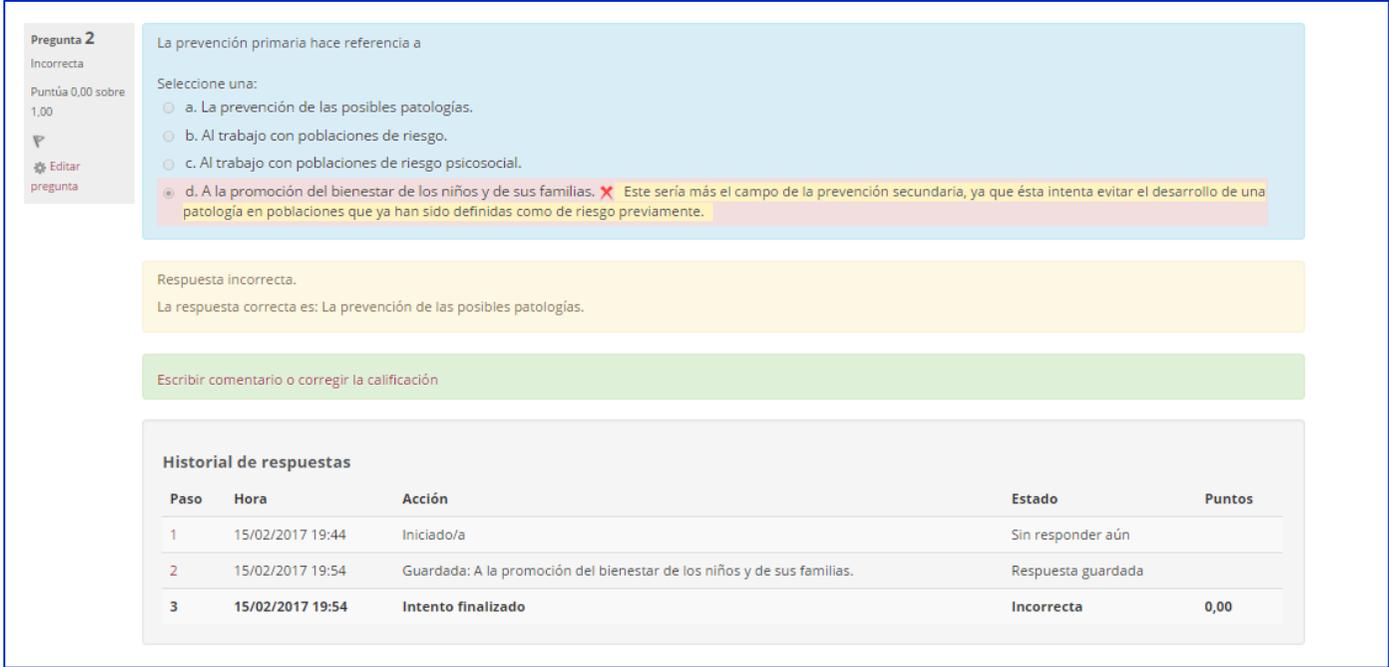
en la asignatura de «Estimulación Temprana», que se imparte en tercer curso en el grado en Terapia Ocupacional. En dicha asignatura se trabajó con la metodología de ABP. Asimismo, se aplicaron distintos tipos de *feedback* a la ejecución de las tareas siguiendo procesos de Tutoría inteligente a través de la LMS (en concreto UBUVirtual) en unos casos automatizadas y en otros personalizadas. En la Tabla 3, se presentan distintos tipos de *feedback* que se realizaron a través de la plataforma.

Tabla 3

Relación entre los conceptos de conocimiento de la cognición, regulación de la cognición y conocimiento metacognitivo

Tarea	Tipo de Feedback	Tipo de interacción
<p><i>Feedback</i> formativo sobre reglas y normativa de resolución del ABP.</p>	<p>Instrucciones en la plataforma</p>	<p>Información inicial</p>
 <p>The screenshot shows a digital library or course page titled 'Aprendizaje Basado en Proyectos'. It lists various documents, rubrics, and collaborative materials. Each item has a plus sign on the left and a set of icons (gear, arrows, eye, person, lock) on the right, indicating interactive options for each resource.</p>		
<p><i>Feedback</i> sobre la ejecución en los cuestionarios</p>	<p><i>Feedback</i> sobre la respuesta correcta y el porqué de la</p>	<p>Interactiva con las ideas de los estudiantes acerca de los conceptos trabajados.</p>

Tarea	Tipo de <i>Feedback</i>	Tipo de interacción																				
de conocimiento o sobre los conceptos trabajados en las unidades temáticas.	misma. <i>Formativo.</i>	 <p>Pregunta 1 Correcta Puntúa 1,00 sobre 1,00 ✖ Editar pregunta</p> <p>La intervención educativa en sujetos con discapacidad se potenció</p> <p>Seleccione una:</p> <ul style="list-style-type: none"> <input checked="" type="radio"/> a. En el siglo XX. ✓ La integración social de las personas con discapacidad se inició en el siglo XX, con medidas legislativas que potenciaban la normalización y la integración social (LISMI, LOGSE...) y en la actualidad estos conceptos se han incluido en el concepto de inclusión social de las personas con discapacidad. <input type="radio"/> b. En el siglo XXVII <input type="radio"/> c. En el siglo XXII <input type="radio"/> d. En el s. XVIII <p>Respuesta correcta La respuesta correcta es: En el siglo XX.</p> <p>Escribir comentario o corregir la calificación</p> <p>Historial de respuestas</p> <table border="1"> <thead> <tr> <th>Paso</th> <th>Hora</th> <th>Acción</th> <th>Estado</th> <th>Puntos</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>15/02/2017 19:44</td> <td>Iniciado/a</td> <td>Sin responder aún</td> <td></td> </tr> <tr> <td>2</td> <td>15/02/2017 19:54</td> <td>Guardada: En el siglo XX.</td> <td>Respuesta guardada</td> <td></td> </tr> <tr> <td>3</td> <td>15/02/2017 19:54</td> <td>Intento finalizado</td> <td>Correcta</td> <td>1,00</td> </tr> </tbody> </table>	Paso	Hora	Acción	Estado	Puntos	1	15/02/2017 19:44	Iniciado/a	Sin responder aún		2	15/02/2017 19:54	Guardada: En el siglo XX.	Respuesta guardada		3	15/02/2017 19:54	Intento finalizado	Correcta	1,00
Paso	Hora	Acción	Estado	Puntos																		
1	15/02/2017 19:44	Iniciado/a	Sin responder aún																			
2	15/02/2017 19:54	Guardada: En el siglo XX.	Respuesta guardada																			
3	15/02/2017 19:54	Intento finalizado	Correcta	1,00																		
<i>Feedback</i> sobre las actividades prácticas sobre los	Respuesta <i>Correcta</i>																					

Tarea	Tipo de Feedback	Tipo de interacción																				
conceptos teóricos.	Respuesta incorrecta	 <p>Pregunta 2 Incorrecta Puntúa 0,00 sobre 1,00 ▼ Editar pregunta</p> <p>La prevención primaria hace referencia a</p> <p>Seleccione una:</p> <ul style="list-style-type: none"> <input type="radio"/> a. La prevención de las posibles patologías. <input type="radio"/> b. Al trabajo con poblaciones de riesgo. <input type="radio"/> c. Al trabajo con poblaciones de riesgo psicosocial. <input checked="" type="radio"/> d. A la promoción del bienestar de los niños y de sus familias. ✘ Este sería más el campo de la prevención secundaria, ya que ésta intenta evitar el desarrollo de una patología en poblaciones que ya han sido definidas como de riesgo previamente. <p>Respuesta incorrecta. La respuesta correcta es: La prevención de las posibles patologías.</p> <p>Escribir comentario o corregir la calificación</p> <p>Historial de respuestas</p> <table border="1"> <thead> <tr> <th>Paso</th> <th>Hora</th> <th>Acción</th> <th>Estado</th> <th>Puntos</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>15/02/2017 19:44</td> <td>Iniciado/a</td> <td>Sin responder aún</td> <td></td> </tr> <tr> <td>2</td> <td>15/02/2017 19:54</td> <td>Guardada: A la promoción del bienestar de los niños y de sus familias.</td> <td>Respuesta guardada</td> <td></td> </tr> <tr> <td>3</td> <td>15/02/2017 19:54</td> <td>Intento finalizado</td> <td>Incorrecta</td> <td>0,00</td> </tr> </tbody> </table>	Paso	Hora	Acción	Estado	Puntos	1	15/02/2017 19:44	Iniciado/a	Sin responder aún		2	15/02/2017 19:54	Guardada: A la promoción del bienestar de los niños y de sus familias.	Respuesta guardada		3	15/02/2017 19:54	Intento finalizado	Incorrecta	0,00
Paso	Hora	Acción	Estado	Puntos																		
1	15/02/2017 19:44	Iniciado/a	Sin responder aún																			
2	15/02/2017 19:54	Guardada: A la promoción del bienestar de los niños y de sus familias.	Respuesta guardada																			
3	15/02/2017 19:54	Intento finalizado	Incorrecta	0,00																		

Tarea

Tipo de
Feedback

Tipo de interacción

Sumativo

UNIVERSIDAD DE BURGOS **campus de Excelencia INTERNACIONAL**

UBU Virtual

MARIA CONSUELO SAIZ MANZANARES Español - Internacional (es)

Comenzado el	miércoles, 15 de febrero de 2017, 19:44
Estado	Finalizado
Finalizado en	miércoles, 15 de febrero de 2017, 19:54
Tiempo empleado	9 minutos 12 segundos
Calificación	7,67 de 10,00 (77%)

Pregunta 1
Correcta
Puntúa 1,00 sobre 1,00

La intervención educativa en sujetos con discapacidad se potenció

Seleccione una:

- a. En el siglo XX. La integración social de las personas con discapacidad se inició en el siglo XX, con medidas legislativas que potenciaban la normalización y la integración social (LISMI, LOGSE...) y en la actualidad estos conceptos se han incluido en el concepto de inclusión social de las personas con discapacidad.
- b. En el siglo XXVII
- c. En el siglo XXII
- d. En el s. XVIII

Respuesta correcta
La respuesta correcta es: En el siglo XX.

Escribir comentario o corregir la calificación

Formativo

Dinámico y personalizado a las acciones de los grupos.

Tarea
Tipo de
Feedback

Tipo de interacción

Seleccionar	Imagen del usuario	Apellidos(s) / Nombre	Dirección de correo	Grupo	Calificación	Última modificación (entrega)	Archivos enviados	Comentarios de la entrega	Última modificación (calificación)	Comentarios de retroalimentación
<input type="checkbox"/>				Grupo 7	Calificación Apto	domingo, 12 de febrero de 2017, 16:57	Práctica 2.pdf Estado de Ephorus: 6%	Comentarios (0)	Jueves, 16 de febrero de 2017, 17:26	Una práctica excelente, muy trabajada. Enhorabuena.
<input type="checkbox"/>				Grupo 8	Calificación Apto	lunes, 13 de febrero de 2017, 17:05	Práctica 2.docx Estado de Ephorus: 4%	Comentarios (0)	Jueves, 16 de febrero de 2017, 17:30	Buena práctica. Para sucesivas ocasiones subir el doc. en formato pdf.
<input type="checkbox"/>				Grupo 9	Calificación Apto	lunes, 13 de febrero de 2017, 17:17	Práctica 2. Grupo 9.docx Estado de Ephorus: 31%	Comentarios (0)	miércoles, 15 de marzo de 2017, 10:41	Dejo revisada la práctica en la wiki, lo vemos en clase la semana que viene.
<input type="checkbox"/>				Grupo 13	Calificación Apto	lunes, 13 de marzo de 2017, 12:17	unidad 2 practica dos.docx Estado de Ephorus: 0%	Comentarios (0)	miércoles, 15 de marzo de 2017, 10:48	Muy buena, práctica enhorabuena. Ya sabes siempre el archivo en pdf.
<input type="checkbox"/>				Grupo 9	Calificación Apto	lunes, 13 de febrero de 2017, 17:17	Práctica 2. Grupo 9.docx Estado de Ephorus: 31%	Comentarios (0)	miércoles, 15 de marzo de 2017, 10:48	Dejo revisada la práctica en la wiki, lo vemos en clase la semana que viene.
<input type="checkbox"/>				Grupo 8	Calificación Apto	lunes, 13 de febrero de 2017, 17:05	Práctica 2.docx Estado de Ephorus: 6%	Comentarios (0)	Jueves, 16 de febrero de 2017, 17:30	Buena práctica. Para sucesivas ocasiones subir el doc. en formato pdf.
<input type="checkbox"/>				Grupo 9	Calificación Apto	lunes, 13 de febrero de 2017, 17:17	Práctica 2. Grupo 9.docx Estado de Ephorus: 31%	Comentarios (0)	miércoles, 15 de marzo de 2017, 10:48	Dejo revisada la práctica en la wiki, lo vemos en clase la semana que viene.

Tarea

Tipo de
Feedback

Tipo de interacción

Formativo

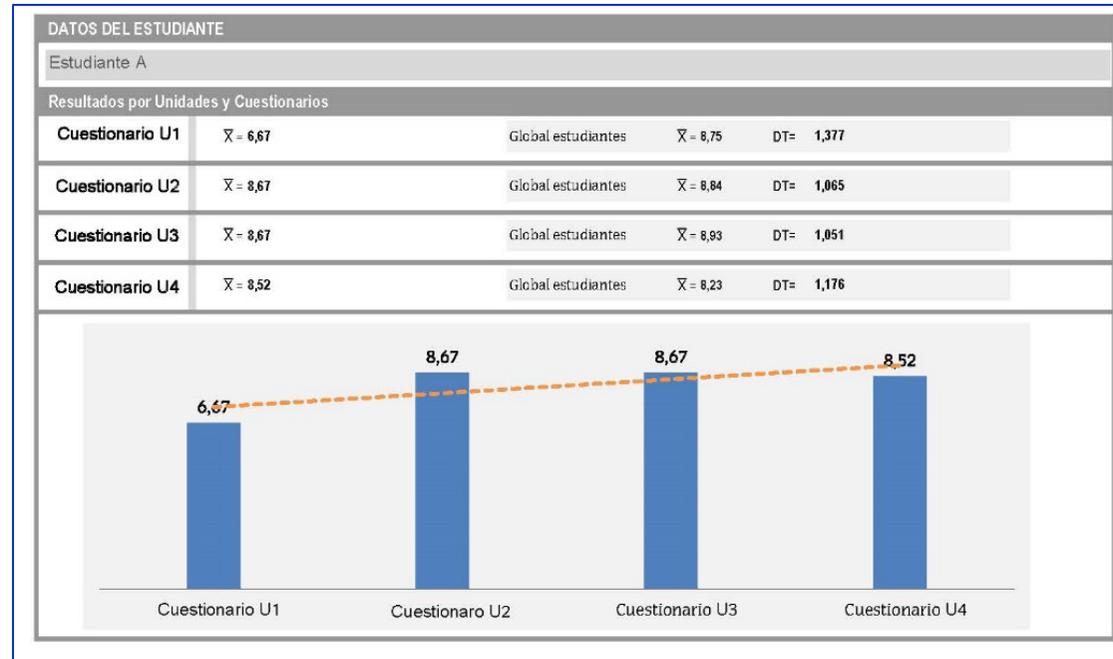
The screenshot shows a Moodle course page for 'Trabajo colaborativo en la elaboración del Proyecto'. The page is in Spanish and features a navigation sidebar on the left with sections for 'Administración' (Administration) and 'Navegación' (Navigation). The main content area displays a list of files and documents, including PDFs and DOCX files, with a 'Ver' (View) button at the bottom. The course is managed by Maria Consuelo Saiz Manzanares. The page title is 'Trabajo colaborativo en la elaboración del Proyecto' and the course ID is 'ESTIMULACIÓN TEMPRANA (5866 #1P)'.

Tarea

Tipo de
Feedback

Tipo de interacción

De
tendencia,
formativo y
sumativo.

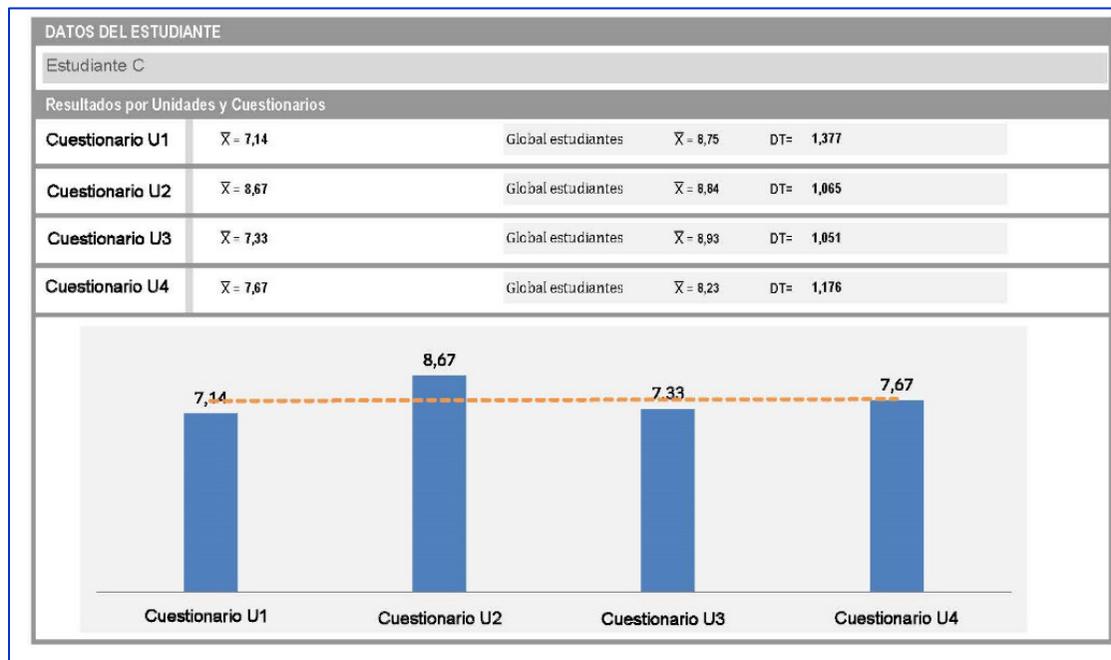


Tarea

Tipo de *Feedback*

Tipo de interacción

De
tendencia,
formativo y
sumativo.

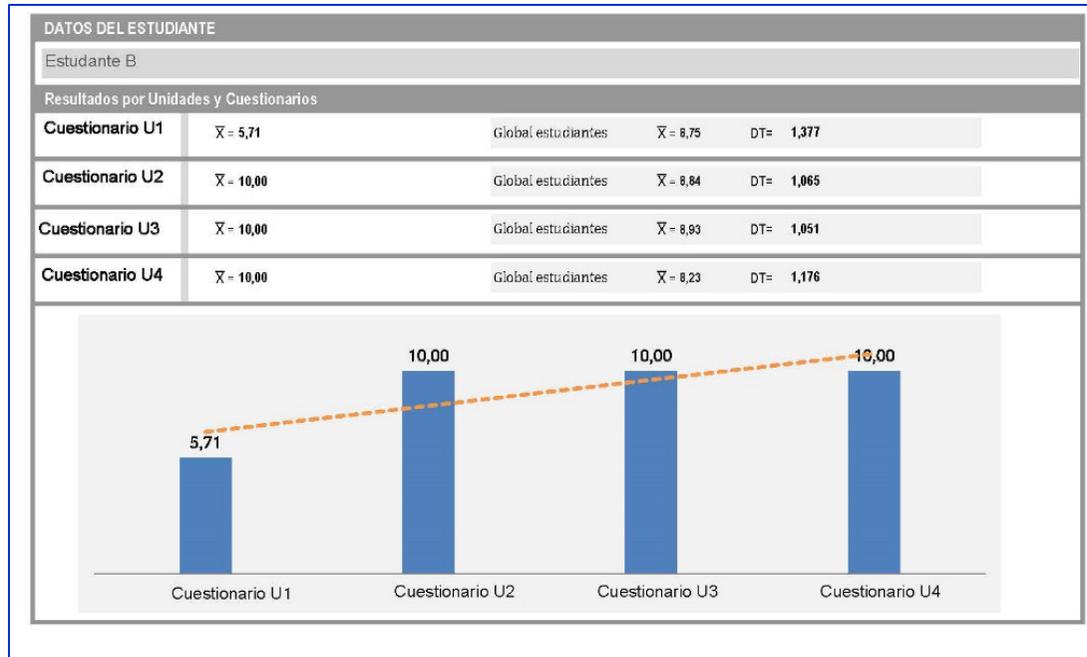


Tarea

Tipo de *Feedback*

Tipo de interacción

De
tendencia,
formativo y
sumativo.



1.7.4. Técnicas de EMD y LA aplicadas al estudio del desarrollo de la asignatura «Estimulación Temprana»

En la asignatura «Estimulación Temprana» se utilizaron diferentes fuentes de datos dirigidas a obtener distinto tipo de información en el proceso de enseñanza-aprendizaje, en la Tabla 4 se pueden consultar y en la Figura 8 se puede ver el resumen del proceso de extracción de datos para su posterior análisis.

Tabla 4

Procedimiento de extracción y análisis de datos para la elaboración de informes en Moodle (UBUVirtual)

Procedimiento de obtención de datos	Extracción de logs Importación a SPSS	Análisis de datos	Informes
Escala de Estrategias metacognitivas (ACRAR).	X	Análisis de los percentiles y de las subestrategias	LA en Moodle
Escala de Apoyo al Procesamiento de la Información (ACRAR).	X	Análisis de los percentiles y de las subestrategias	LA en Moodle
Escala de conocimientos previos sobre atención temprana.	X	Medias	LA en Moodle
Escala de Evaluación final del desarrollo de la asignatura.	X	Medias	LA en Moodle
Encuesta sobre el desarrollo de cada unidad temática.	X	Medias	LA en Moodle
Cuestionario de evaluación por unidad temática.	X	Análisis cuantitativo, individual y del grupo.	LA en Moodle
Cuestionario de evaluación de los vídeos complementarios en la modalidad FC.	X	Medias	LA en Moodle

Nota. LA = *Learning Analytics*

Un ejemplo de obtención de informes desde LA en UBUVirtual se puede consultar en la Figura 8.

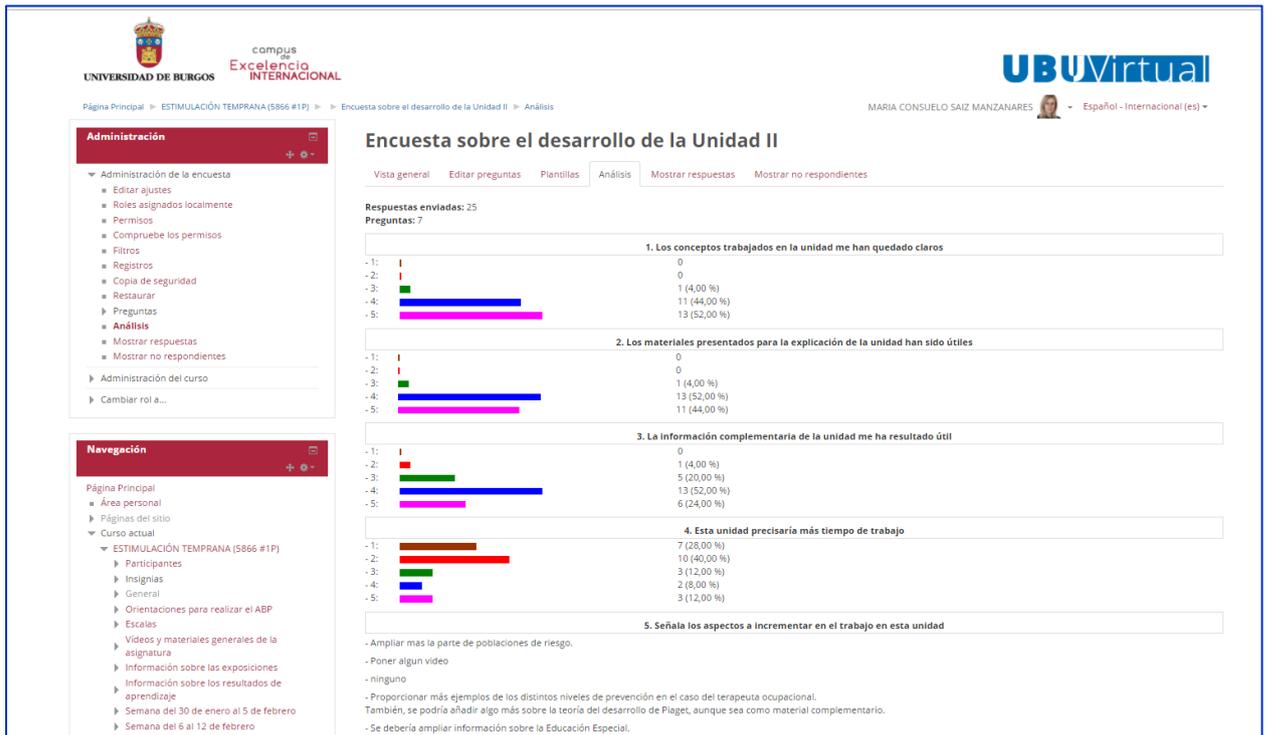


Figura 8

Análisis de los resultados en la encuesta sobre el desarrollo de la Unidad Temática II en la asignatura «Estimulación Temprana»

También, se pueden extraer los *logs* de información de la plataforma UBUVirtual para posteriormente poder analizarlos en paquetes estadísticos como SPSS, R o Matlab o de minería de datos como WEKA, R o bibliotecas de Python. Dichos programas permiten realizar análisis estadísticos más complejos. El proceso de extracción y posterior vinculación a dichos análisis se puede consultar en la Figura 9.

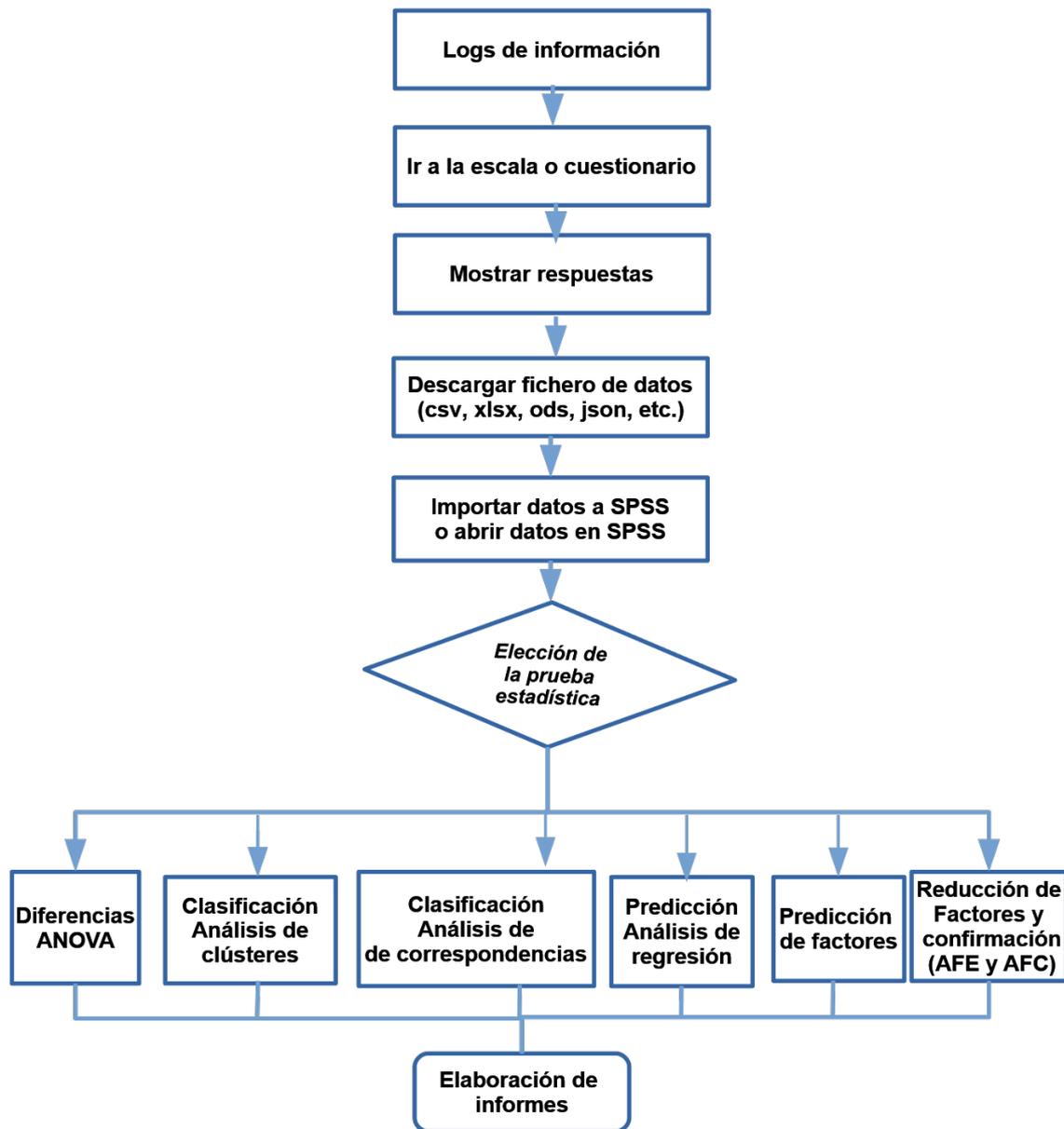


Figura 9

Secuencia de pasos para la extracción de datos en UBUVirtual y análisis de datos en SPSS

A continuación, se presenta un ejemplo del procedimiento de ejecución desde la secuenciación de pasos (ver de la Figura 10 a la Figura 14).



Figura 10

Paso 1 de 5: elección de la Escala a analizar

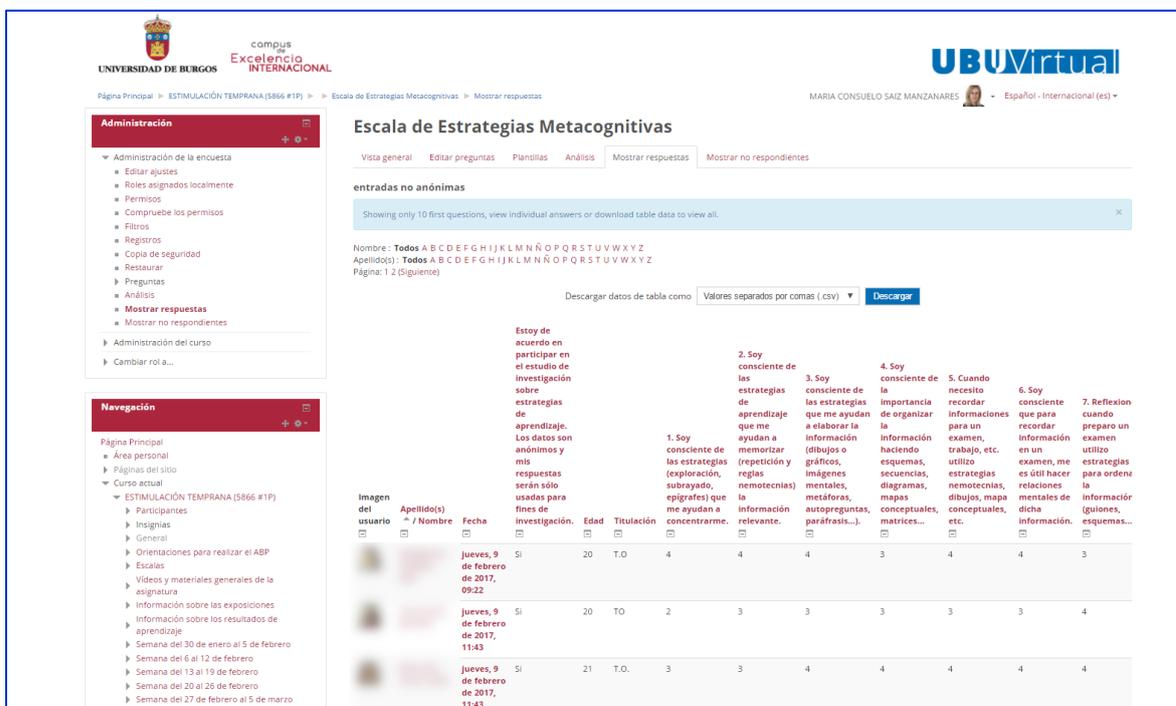


Figura 11

Paso 2 de 5: Mostrar respuestas y elegir del tipo de fichero para la extracción de los datos

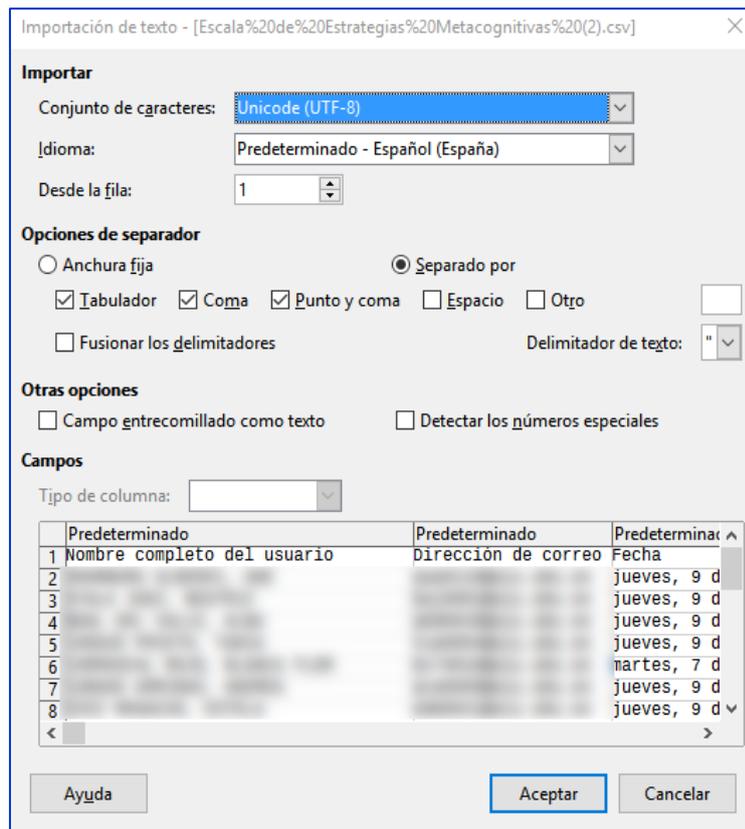


Figura 12

Extracción de datos en un fichero CSV. Paso 2 y 3 de 5

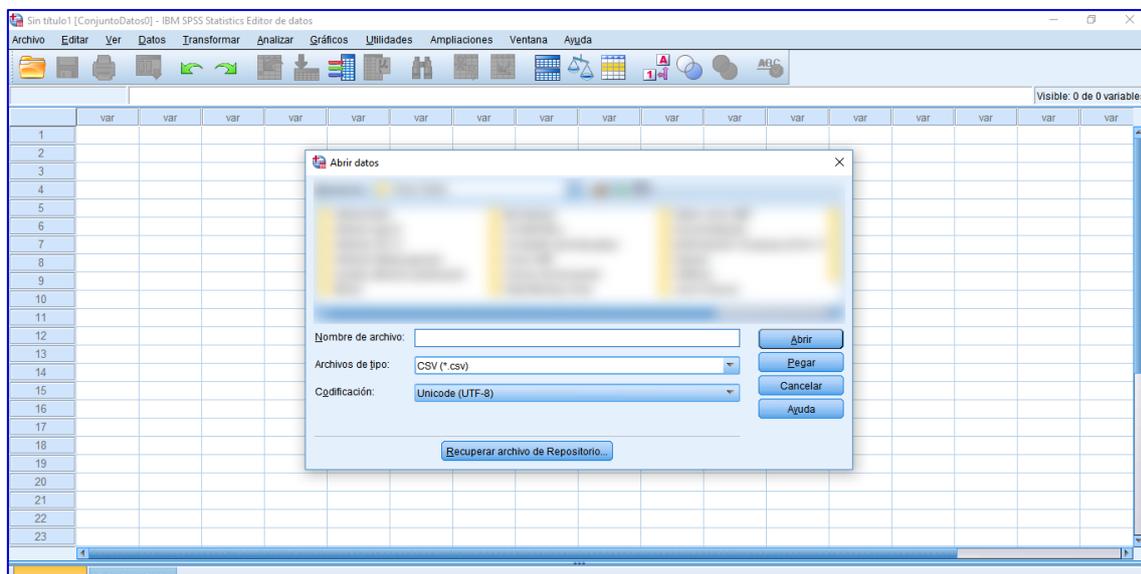


Figura 13

Paso 4 de 5: abrir el fichero CSV en el programa SPSS

	Fecha	Estoyd	Edad	Titulac...	@1.Soyconsciente delasestrategiasde loraciónsubraya...	@2.Soyconsciente delasestrategiasde prendizajequem...	@3.Soyconsciente delasestrategiasde meayudanaelabor...	@4.Soyconsciente delaimportanciaeoi ganizarlainforma...	@5.Cuandonecesit orecordarinformacio nesparanexame...	@6.Soycons quepararecor rmaciónenun...
1	jueves, 9 de febrero de 2017, 09:22	Si	20	T.O	4	4	4	3	4	
2	jueves, 9 de febrero de 2017, 11:43	Si	20	TO	2	3	3	3	3	
3	jueves, 9 de febrero de 2017, 11:43	Si	21	T.O	3	3	4	4	4	
4	jueves, 9 de febrero de 2017, 11:41	Si	20	TO	4	4	4	4	4	
5	martes, 7 de febrero de 2017, 10:56	Si	20	TO	3	3	3	4	4	
6	jueves, 9 de febrero de 2017, 12:39	Si	20	TO	3	4	4	4	4	
7	jueves, 9 de febrero de 2017, 11:48	Si	20	T.O	4	4	3	2	2	
8	jueves, 9 de febrero de 2017, 11:41	Si	25	TO	3	3	3	4	4	
9	sábado, 11 de febrero de 2017, 14:01	Si	22	T.O	4	4	3	3	4	
10	jueves, 9 de febrero de 2017, 17:32	Si	23	Grado	4	3	4	4	3	
11	jueves, 9 de febrero de 2017, 11:43	Si	27	T.O	3	3	3	3	3	
12	martes, 7 de febrero de 2017, 10:55	Si	25	TO	4	4	4	4	4	
13	martes, 7 de febrero de 2017, 10:50	Si	23	TO	3	3	3	3	4	
14	domingo, 12 de febrero de 2017, 16:41	Si	21	T.O	4	3	4	3	3	
15	viernes, 31 de marzo de 2017, 09:09	Si	32	T.O	4	2	3	4	4	
16	viernes, 10 de febrero de 2017, 14:38	Si	20	T.O	4	4	4	4	4	
17	jueves, 9 de febrero de 2017, 12:53	Si	21	TO	4	3	3	4	3	
18	domingo, 12 de febrero de 2017, 19:06	Si	21		3	4	4	4	3	
19	miércoles, 22 de marzo de 2017, 13:18	Si	21	TO	3	3	3	3	4	
20	sábado, 11 de febrero de 2017, 13:44	Si	23		4	3	3	3	3	
21	jueves, 9 de febrero de 2017, 11:40	Si	20	T.O	4	3	4	3	3	
22	miércoles, 22 de febrero de 2017, 13:07	Si	20	T.O	3	3	2	2	2	

Figura 14

Paso 5 de 5. Transformación para su análisis de los datos en formato CSV a un fichero SPSS

Un diagrama con los pasos de extracción y de análisis de datos se puede observar en la Figura 15.

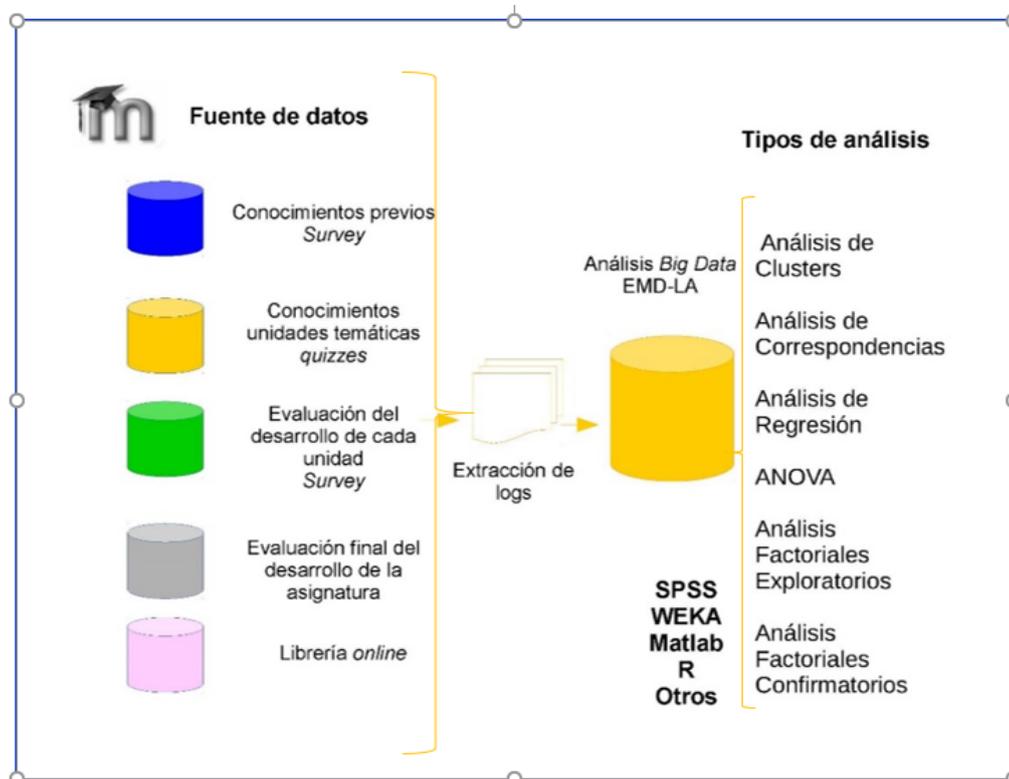


Figura 15

Esquema para el procedimiento de ejecución del análisis de datos en Moodle

Atendiendo a lo visto anteriormente, los sistemas de tutoría inteligente permiten adaptar y secuenciar los contenidos al ritmo de aprendizaje del estudiante personalizando el *feedback* (Verdú, Regueras, Gal, de Castro, Verdú, y Kohen-Vacs, 2016). Actualmente, estas investigaciones van dirigidas a incluir la utilización de mecanismos de extensión como son los *blocks* (almacenamiento de información en bloques en la nube). Estos son *plugins*² que provienen de un *box* (herramientas virtuales que promueven y desarrollan el almacenamiento de la información en la nube) con contenido contextual. Los bloques aparecen alrededor de los bordes de la página o escondidos como *popups* (ventanas emergentes) de Moodle (ventanas emergentes en Moodle). Dichos accesorios enriquecen el contenido de la página e incluyen un código personalizado al generar una página en el servidor. Cada vez que un estudiante realiza una actividad en Moodle se conecta con el tutor inteligente que en función de la acción del alumno le devuelve un comentario, *feedback* o sugerencia. La incorporación de estas ayudas, sistemas de tutoría inteligente, facilita el desarrollo del aprendizaje de los estudiantes. La razón es que potencia el aprendizaje personalizado y respeta el ritmo de aprendizaje del alumno. Todo ello, permite al estudiante desarrollar su aprendizaje en los espacios elegidos por él, por lo que previsiblemente incrementará su motivación. Se basa en el lenguaje OWL (lenguaje de Ontologías Web) y se ha diseñado para ser usado en aplicaciones que precisan procesar el contenido de la información y no sólo representarla (Verdú et al., 2016). Los Sistemas de Tutorización Inteligente, pues, pueden ayudar en la orientación en la navegación, en el orden de consulta del contenido, en el desarrollo del proceso de aprendizaje, en el desarrollo del *feedback*, en los procesos de auto-evaluación y en el uso de los contenidos hipermedia entre otros procesos (Kacprzyk y Jain, 2017). Si bien, en la actualidad Moodle no proporciona ninguna técnica compleja de minería de datos relacionada con los sistemas de predicción y *clustering*, solo realiza análisis sencillos de medias y de la fiabilidad de los cuestionarios de conocimientos, en este último caso también proporciona datos sobre la distribución de las puntuaciones de los participantes (Luna, Castro, y Romero, 2017). Por ello, sería interesante el diseño de una estructura de extracción de los *logs* de información relativa a la navegación de los estudiantes en las distintas actividades de aprendizaje, así como en los procedimientos de evaluación que permita un estudio analítico de su ritmo y trayectoria de aprendizaje. Este estudio servirá tanto para

² Es una aplicación o programa informático que se relaciona con otro para agregar una función específica, esta aplicación adicional es incluida por la principal y actúa a través de un interfaz de programación de aplicaciones.

efectuar el oportuno *feedback* al estudiante como para rediseñar dicho Sistema de Tutorización Inteligente, un ejemplo para la lógica de inserción de las técnicas en Moodle se presenta en la Figura 16.

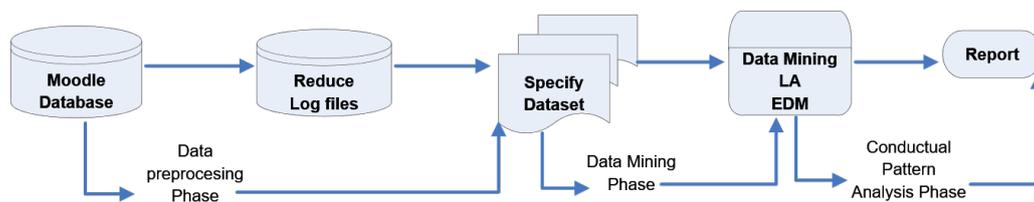


Figura 16

Esquema de las fases en la arquitectura del sistema inteligente en Moodle, adaptado del modelo de Romero y Ventura (2007)

1.7.5. Recursos hipermedia en la personalización del aprendizaje

Los recursos hipermedia se ubican en los entornos de e-aprendizaje. Estos recursos combinan hipertextos, *hyperlinks* y materiales audio-visuales para el acceso a la información objeto de aprendizaje (Kacprzyk y Jain, 2017). La orientación al estudiante se puede hacer por dos vías, de una forma individual o grupal, o ambas. Ya que se puede orientar los caminos de aprendizaje del sujeto y los del grupo. El profesor diseña los contenidos educativos desde una orientación de secuenciación de objetivos los contenidos y criterios de evaluación implementando su desarrollo desde sistemas hipermedia. El objetivo final es el desarrollo eficaz de las competencias. Un ejemplo de este uso es la elaboración de vídeos interactivos en experiencias *Flipped Classroom* (FL) dentro de entornos *Flipped Learning* (FL). Estos vídeos facilitan la integración de la información presentada por varios canales (audio, vídeo, animación, gráficos...), además de la navegación individualizada respetando el ritmo de aprendizaje de cada estudiante. Los recursos hipermedia pueden ser variados: vídeos con cuestionarios de comprensión, lecturas, sistemas de tutoría inteligente y actividades interactivas en clase (Johnston, 2017). No obstante, el desarrollo del proceso de aprendizaje desde una estructura de tutorización inteligente a través del uso de las plataformas interactivas implica un profundo y sistemático proceso de planificación por parte del docente (McGivney-Burelle y Xue, 2013).

1. Estructuración del tiempo.
2. Elaboración de los procedimientos de evaluación.
3. Inclusión de técnicas de ABP.
4. Diseño y desarrollo del *Feedback* F2F y virtual.
5. Diseño de procedimientos para la evaluación continua de la práctica docente por parte de los alumnos.

Asimismo, el uso de los vídeos interactivos ayuda al *Self Regulated Learning* (SRL) (Moos y Blonde, 2016). Además, la inclusión de cuestionarios con preguntas de comprensión y con *feedback* de orientación después de la ejecución de cada ítem facilita la comprensión y la autorregulación del aprendizaje y por ende la construcción del conocimiento. Los vídeos se pueden diseñar desde la regulación de las estrategias metacognitivas. Para lograrlo un elemento importante en el diseño de la asignatura es el utilizar la taxonomía de Bloom (2008) para elaborar las distintas actividades de aprendizaje que permitan una construcción interactiva del conocimiento (Schneider y Blikstein, 2016) y el desarrollo de estrategias cognitivas y metacognitivas. En la Figura 17 se puede observar la estructura de análisis de las competencias del estudiante en las distintas tareas de aprendizaje según Bloom (2008) y en el Figura 18 un diseño de tutorización inteligente desde la utilización de los recursos hipermedia.

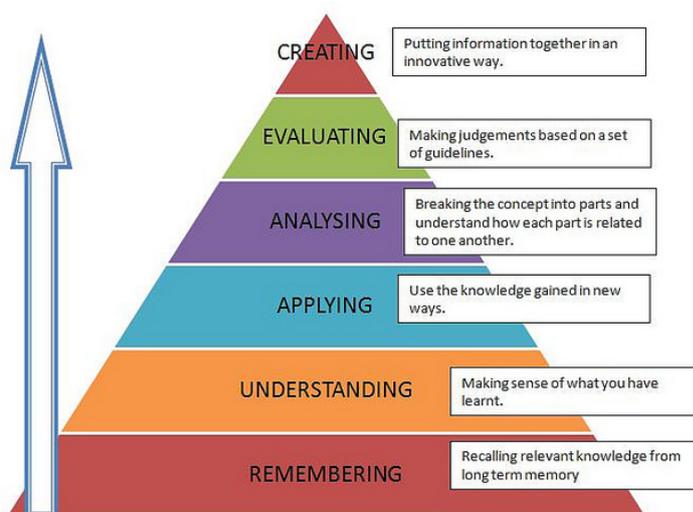


Figura 17

Habilidades competenciales en la Taxonomía de Bloom (2008)

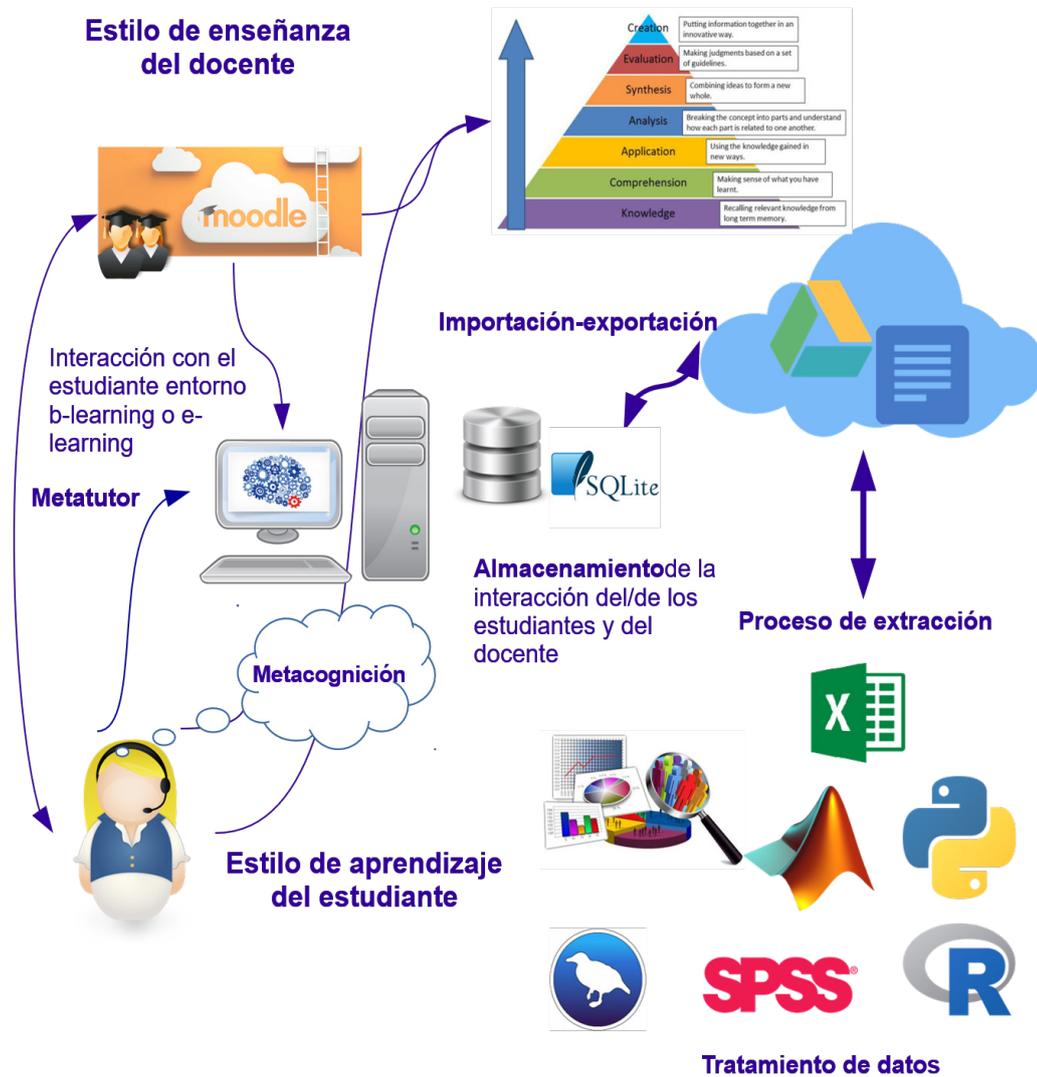


Figura 18

Proceso de tutorización personalizada desde sistemas de tutoría inteligente con componentes hipertexto fuente propia

De una forma sencilla en la Figura 19 y Figura 20 se presenta un diseño de tutorización inteligente en la asignatura de «Estimulación Temprana».

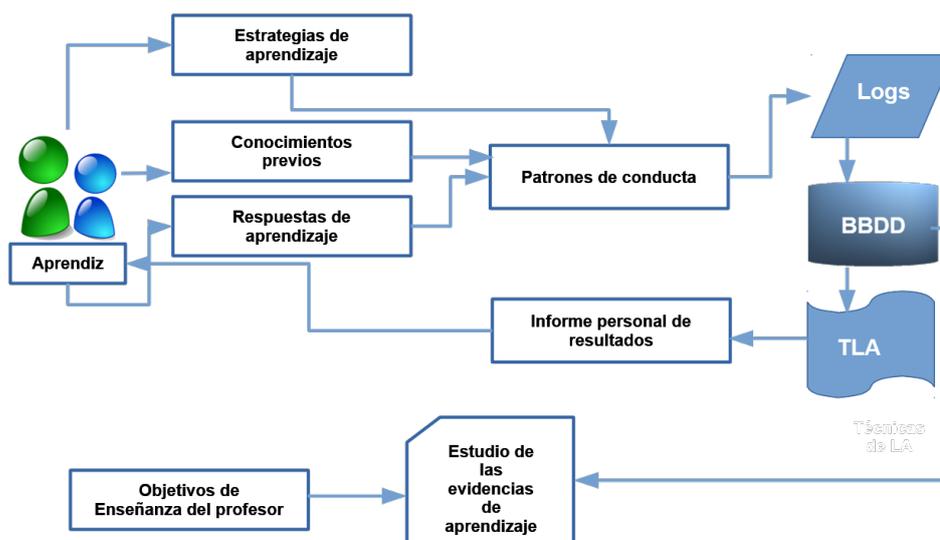


Figura 19

Diseño de un Sistema de Arquitectura de Tutoría Inteligente en la plataforma Moodle

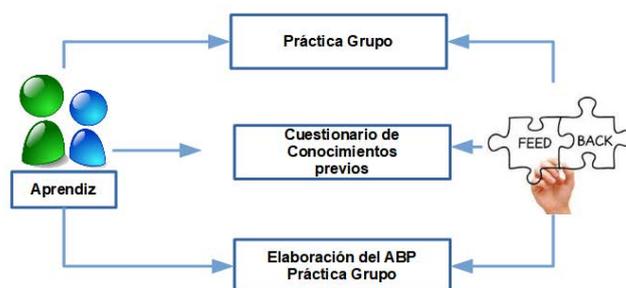


Figura 20

Diseño de un Sistema de Arquitectura de Tutoría Inteligente en la plataforma Moodle para el diseño de la asignatura de «Estimulación Temprana»

1.8. Técnicas Data Mining (DM) y de Pattern Mining (PM) en el entorno Educativo

En este apartado se describirán pormenorizadamente los procesos de extracción y de análisis de los datos que se desprenden de la interactividad en la plataforma Moodle. Dicha interacción depende en un porcentaje alto del diseño que de la docencia haga el profesor, o

bien marque la institución como sello de identidad. Partiendo del supuesto de que se ha planificado un entorno de interacción basado en los recursos de *feedback* orientado a procesos, utilización de la metodología de aprendizaje basado en proyectos, uso de recursos hipermedia (uso de metodologías activas). Seguidamente, se va a señalar cómo se pueden extraer, depurar y procesar los datos registrados en el entorno LMS basado en Moodle, que en este caso es la plataforma UBUVirtual. Previamente a la extracción de los datos, el profesor investigador tendrá que intentar dar respuesta a distintas preguntas de investigación (García, Luengo, y Herrera, 2015) y que están relacionadas con los siguientes aspectos:

- *Especificación del problema*: se realiza teniendo en cuenta estudios anteriores que revelan el estado del Arte en esa disciplina.
- *Comprensión del problema*: esta fase incluye la comprensión de la selección de datos dirigidos hacia el enfoque de conocimiento concreto con el fin de alcanzar un alto grado de fiabilidad.
- *Preprocesamiento de los datos*: este estadio contempla operaciones de depuración de las bases de datos. Esta depuración se realiza eliminando ruido o datos inconsistentes para el objeto de la investigación, integrando datos que vienen de varias fuentes, transformando datos. La transformación en muchos casos significará una reducción y en otros una adaptación a las Técnicas de Minería de Datos que se van a aplicar teniendo en cuenta los softwares donde se van a introducir las bases de datos para su posterior procesamiento.
- *Aplicación de Técnicas de Minería de Datos*: son técnicas de análisis de datos que pueden ser supervisadas o no supervisadas. En dichas técnicas se emplean distintos algoritmos para distintos fines u objetivos. Además, se evalúa el rendimiento de los distintos algoritmos de minería de datos para obtener un valor cuantificable de la calidad del modelo predictivo construido. El objetivo final es la extracción de patrones de comportamiento de los datos que servirán para comprobar las hipótesis de investigación.
- *Aplicación de técnicas de evaluación*: Estas estiman e interpretan los patrones hallados.
- *Explotación de los resultados*: Está directamente relacionado con la visualización e interpretación de los datos.

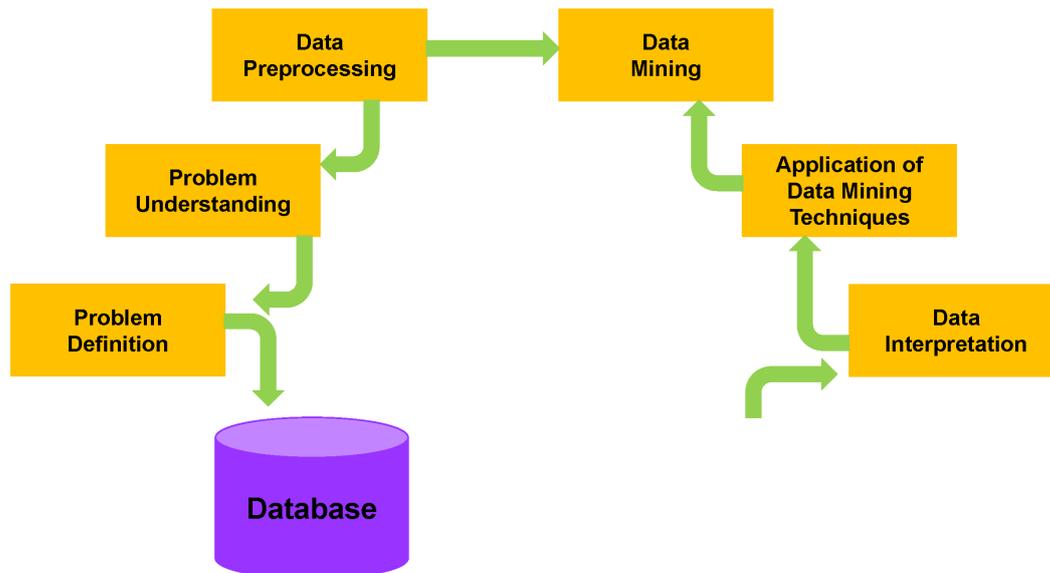


Figura 21

Imagen adaptada del original de García, Luengo, y Herrera (2015) p. 3 sobre *Knowledge Discovery in Databases (KDD)*

En este proceso es especialmente importante la fase de preparación de los datos o de preprocesamiento, siguiendo a García, Luengo, y Herrera (2015, p.10-13) se pueden distinguir las siguientes fases:

- *Fase de depuración de los datos (data cleaning)*. Las bases de datos requieren un análisis de los valores y una detección, en su caso, de los valores que faltan así como un análisis de toda aquella información que no es relevante para el propósito de la investigación.
- *Fase de Transformación de los datos (data transformation)*. Los datos tienen que prepararse para poder aplicar sobre ellos distintos algoritmos de Minería de Datos. Para ello, el investigador tiene que saber cuál es el objetivo de su investigación y, en función de ello, qué datos necesita. Además, de conocer las limitaciones de los distintos algoritmos en cuanto al tipo de datos con el que puede trabajar.
- *Fase de integración de los datos (data integration)*. Implica la fusión de datos desde múltiples bases. Este proceso debe realizarse con mucha precisión, con el fin de evitar redundancias e inconsistencias en el resultado final. Las operaciones más usuales son: identificación y unificación de variables y detección de conflictos entre los valores que proceden de distintas fuentes.

- *Fase de normalización de los datos (data normalization)*. La unidad de medida utilizada puede afectar al análisis de datos. Por ello, todos los atributos³ deben ser expresados en las mismas unidades de medida para poder realizar comparaciones entre ellos. La normalización de los datos intenta dar a todos los atributos el mismo peso.
- *Fase de análisis de los valores perdidos (missing data)*. El objetivo de esta fase es eliminar los valores perdidos, se considera que es mejor realizar estimaciones por distintos procedimientos que dejar los valores en blanco. Se pueden utilizar distintas técnicas, por ejemplo: *media de series* (sustituye los valores perdidos con la media de la serie completa), *media de puntos cercanos* (sustituye los valores perdidos por la mediana de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos, por encima y por debajo del valor perdido, utilizados para calcular la mediana), *interpolación lineal* (sustituye los valores perdidos utilizando una interpolación lineal. Se utilizan para la interpolación el último valor válido antes del valor perdido y el primer valor válido después del valor perdido. Si el primer o el último caso de la serie tiene un valor perdido, el valor perdido no se sustituye), y *tendencia línea en el punto* (reemplaza los valores perdidos de la serie por la tendencia lineal en ese punto. Se hace una regresión de la serie existente sobre una variable índice escalada de 1 a n . Los valores perdidos se sustituyen por sus valores pronosticados). Si bien, dentro de las Técnicas de Minería de Datos no hay valores primero y último. Por lo que para predecir el valor perdido se utiliza un modelo de Regresión Lineal entrenado con todos los datos menos el desconocido.
- *Fase de identificación del ruido (noise identification)*. Este es paso se entiende como un paso de limpieza de datos, también es conocido como una fase de suavizado en la transformación de datos. Su objetivo principal es el de detectar errores aleatorios o variaciones en una variable medida. Es preferible la detección del ruido en lugar su eliminación. Una vez detectado se puede aplicar un proceso basado en la corrección que podría implicar algún tipo de operación subyacente.

³ Un atributo se entiende como la categorización que hacen las Técnicas de Machine Learning desde la fijación de valores y la predefinición de las características o de los atributos (Witten y Eibe, 2005).

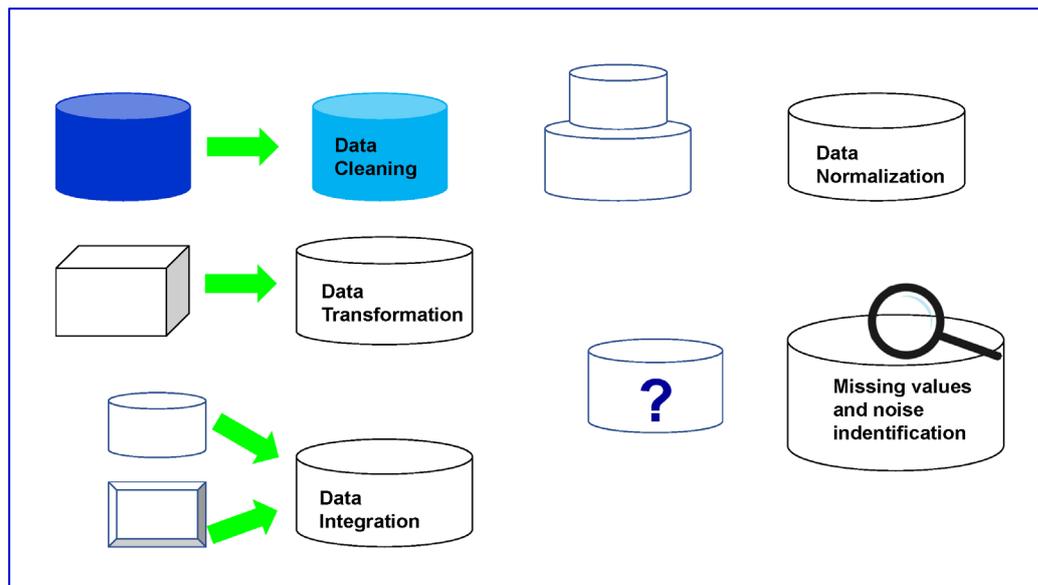


Figura 22

Imagen adaptada del original de García, Luengo, y Herrera (2015) preparación de los datos p. 12

Además, dentro del preprocesamiento está la reducción de datos. Esta incluye un conjunto de técnicas que permite una representación reducida de los datos originales. Entre ellas destacan la selección de las características, la selección de instancias⁴, el *data squashing*⁵ y la discretización (Liu, Hussain, Tan, y Dash, 2002). En concreto, la selección de instancias se puede realizar a través de distintos métodos: muestreo, Boosting⁶, selección de prototipos o de aprendizaje basado en instancias⁷ y aprendizaje activo. Dentro de la selección de prototipos se pueden distinguir: la selección basada en reglas de vecino más cercano, selección basada en la eliminación ordenada, la utilización de algoritmos evolutivos y un muestreo aleatorio (Reinartz, 2002) ver Figura 23 y Figura 24.

⁴ Una instancia es el conjunto de atributos que forman las entidades para las que se dispone de datos de entrada, son las instancias las que hay que clasificar, asociar o agrupar (Witten y Eibe, 2005).

⁵ Se define como la construcción de una base de datos más pequeña que proporciona aproximadamente los mismos resultados que con la base larga (DuMouchel, 2002).

⁶ Boosting es un meta-algoritmo de aprendizaje automático que reduce el sesgo y varianza en un contexto de aprendizaje supervisado. Boosting está basado en el cuestionamiento planteado por Kearns y Valiant (1989) sobre si un conjunto de clasificadores débiles pudiera crear un clasificador robusto. Un clasificador débil está definido para ser un clasificador el cual está solo débilmente correlacionado con la clasificación correcta (el mismo clasifica mejor que un clasificador aleatorio). En contraste, un clasificador robusto es un clasificador que tiene un mejor desempeño que el de un clasificador débil, ya que sus clasificaciones se aproximan más a las verdaderas clases.

⁷ En estos métodos el conjunto de entrenamiento se almacena en su totalidad. Seguidamente, cuando se quiere clasificar una nueva instancia, se usa una función de distancia (típicamente distancia euclídea) para determinar cuál o cuáles son las instancias de entrenamiento más próximas, la clase asignada a la nueva instancia es la de la instancia más próxima, o se determina por votación entre las instancias más próximas. Algunas variantes de este enfoque inicial utilizan una ponderación de los votos que depende de la distancia de la nueva instancia a las instancias en el conjunto de entrenamiento. Para aumentar la rapidez del método, se utilizan algoritmos de búsqueda eficientes, como los KD-Trees. Para evitar los altos requisitos de almacenamiento, cuando el conjunto de entrenamiento es grande, se utilizan técnicas de selección de instancias, que buscan conseguir capacidades predictivas similares a almacenar todo el conjunto de entrenamiento, pero guardando sólo un subconjunto significativo (Olvera-López, Carrasco-Ochoa, Martínez-Trinidad, y Kittler, 2010; Wilson, 1972). Esta selección de instancias ayuda también a aminorar el otro problema de estos métodos, la baja tolerancia al ruido. Aunque es una técnica inicialmente propuesta para resolver problemas de clasificación, ha sido adaptada con éxito en la solución de problemas de regresión. Asimismo, se puede usar boosting para seleccionar instancias se hace un muestreo ponderado en el que tienen más peso las instancias que han resultado difíciles de clasificar correctamente para un modelo anterior (Arnaiz-González, Díez-Pastor, Rodríguez, y García-Osorio, 2016a; 2016b).

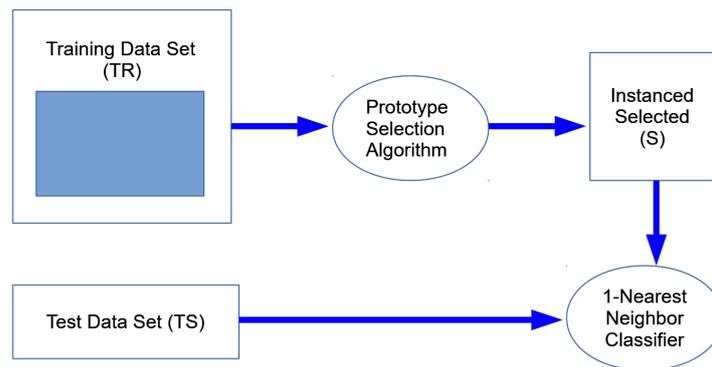


Figura 23

Proceso de selección de instancias. Tomado de Herrera, Riquelme, y Ruíz (2004). Reunión nacional de Data Mining y Machine Learning (Madrid)

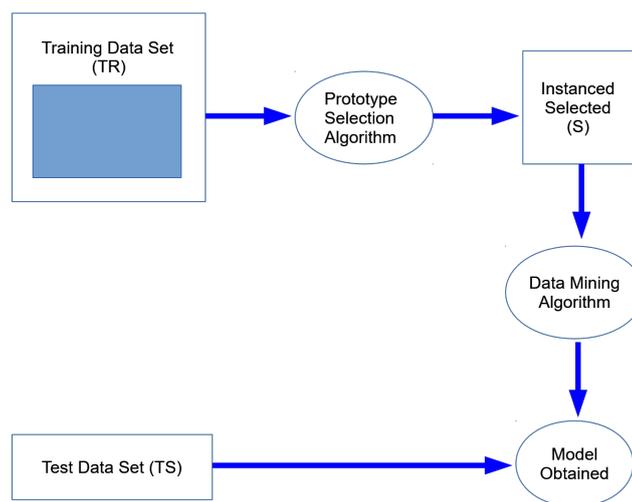


Figura 24

Proceso de selección de instancias. Tomado de Herrera, Riquelme, y Ruíz (2004). Reunión nacional de Data Mining y Machine Learning (Madrid)

La selección de prototipos permite que el conjunto de datos original se transforme en un conjunto de «prototipos», instancias artificiales, no presentes en el conjunto de datos original. En una base de datos las instancias se corresponderían con las filas y los atributos con las columnas (Witten y Eibe, 2005). Dicha selección, mantiene la capacidad predictiva del conjunto original. Trasladado al plano educativo las instancias se corresponderían con los sujetos y los atributos (variables independientes) con el conjunto de valores con los que

se caracteriza a éstos, dichos valores se pueden medir en una escala nominal, ordinal y de intervalos.

En el punto de tratamiento de los datos, se pueden aplicar técnicas estadísticas tradicionales y/o técnicas de minería de datos García, Luengo, y Herrera (2015) ver Figura 25.

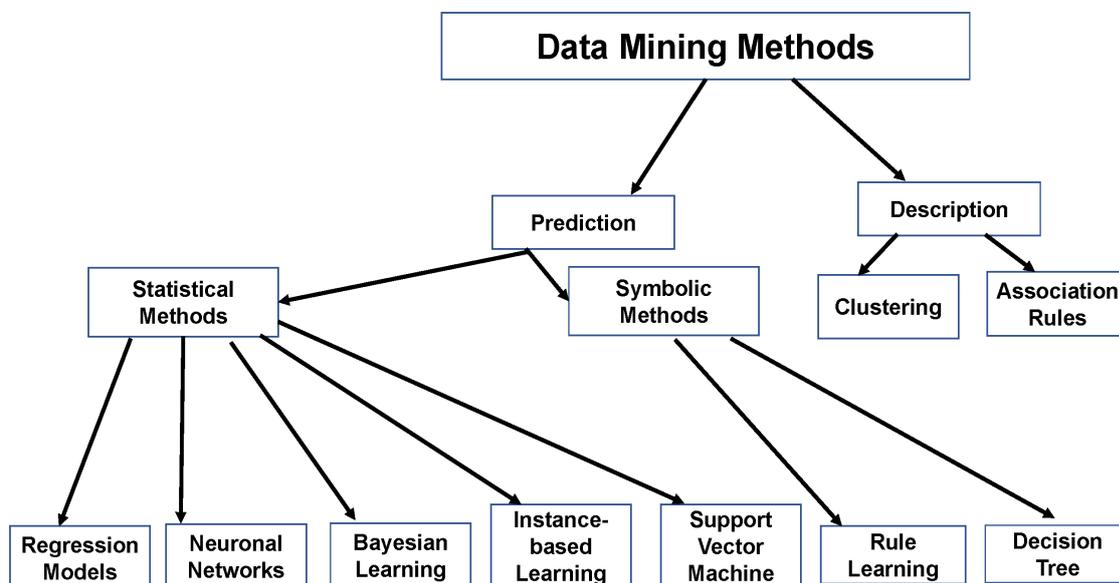


Figura 25

Imagen de elaboración propia sobre Data Mining Methods de García, Luengo, y Herrera (2015) p. 4

Estos autores diferencian entre técnicas de predicción y técnicas de descripción. Dentro de las primeras a su vez distinguen entre métodos estadísticos y métodos simbólicos. Dentro de los primeros se incluyen los modelos de regresión, las redes neuronales, las técnicas bayesianas, el aprendizaje basado en instancias y *Support Vector Machine*. Asimismo, dentro de los métodos simbólicos sitúan el aprendizaje de reglas⁸ y el Árbol de Decisión. Respecto de las técnicas de descripción incluyen *Clustering* y Reglas de Asociación.

⁸ Estas técnicas buscan una regla que explique parte de los datos, separe estos ejemplos y recursivamente separe el resto ejemplos, obteniéndose al final un conjunto de reglas que permite el particionado completo del conjunto (Fürnkranz, Gamberger, y Lavrač, 2012) En términos generales, esta técnica requiere datos nominales o discretizados (aunque esta tarea con frecuencia está implícita en el algoritmo) y procura una selección de atributos interesantes sobre los datos. Ejemplos de algoritmos para encontrar las reglas que permiten este particionado del conjunto de datos son: RIPPER y FURIA. Cuando las reglas no se utilizan para particionar el conjunto de datos, sino para encontrar relaciones interesantes entre los atributos que definen el conjunto, se habla de aprendizaje de reglas de asociación. En origen, las reglas de asociación se utilizaron para analizar la cesta de la compra de los clientes de grandes superficies y para identificar aquellos productos cuya compra se hacía emparejada. Esta información podía utilizarse para mejorar la distribución de los productos en el supermercado y para decidir la política de ofertas de precios. También se han utilizado para el análisis de los accesos Web. Algunos de los algoritmos son: *Apriori*, ECLAT y *FP-growth*.

Otro posible esquema de distribución de las distintas Técnicas de aprendizaje es el que diferencia entre técnicas de *Machine Learning* de aprendizaje supervisado (*supervised learning*) y de aprendizaje no supervisado (*unsupervised learning*). Dentro de las primeras se encuentran las técnicas de clasificación y dentro de ellas se incluyen *Support Vector Machine*, *Discriminant Analysis*, *Naïve Bayes*⁹ y *Nearest Neighbor*. Y las Técnicas de Regresión y dentro de ellas se diferencian las técnicas de Regresión Lineal (*Linear Regression*), *Support Vector Machine*, *Decision Trees* y *Neural Networks*. Asimismo, dentro de las técnicas de aprendizaje no supervisadas se incluyen las Técnicas de *Clustering* y dentro de ellas las técnicas de *k-means*, *k-medoids*, *S-Means*, *Gaussian Mixture* y *Neural Networks* y algunos tipos de redes neuronales como las Redes de función de base radial (RBF) (ver Figura 26).

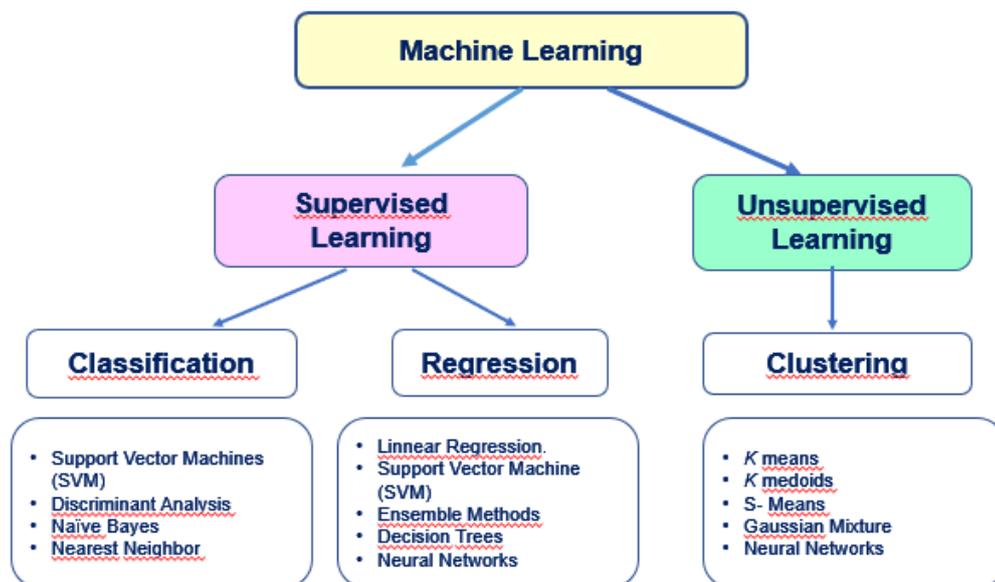


Figura 26

Clasificación de las Técnicas de Machine Learning

⁹ Se basan en el teorema de Bayes, la técnica más utilizada dentro de ellas es el método Naïve Bayes. Dicho método asume que el efecto de un valor de atributo de una clase dada es independiente de los valores de otros atributos. Estos algoritmos sólo se utilizan con atributos categóricos, debido a que el cálculo de probabilidad se hace computacionalmente intratable para dominios no discretos. Además, la suposición de independencia entre los atributos es muy sensible a la redundancia y a la utilidad de algunos de los atributos y ejemplos de los datos. Estas técnicas no pueden trabajar con valores perdidos. Además, de Naïve Bayes, también hay otros modelos complejos basados en estructuras de dependencia como las Redes Bayesianas, también conocidas como modelos gráficos probabilísticos (Koller y Friedman, 2009). Estos modelos son ejemplos de desarrollos dentro de la estadística bayesiana, que es una rama de la estadística en gran auge en la actualidad (Puga, Krzywinski, y Altman, 2015).

Siguiendo a García, Luengo, y Herrera (2015), el objetivo de las técnicas de aprendizaje supervisado es el descubrimiento de las relaciones entre los atributos de entrada (a veces llamados variables o características) y un atributo de destino (a veces referido como clase). La relación que se busca se representa en una estructura llamada modelo. Este se describe como datos ocultos que se explicitan en una predicción del valor del atributo de destino, cuando los valores de los atributos de entrada son conocidos. El aprendizaje supervisado se utiliza, entre otros, en campos de las ciencias de la salud y de la ingeniería. Frecuentemente, se emplea para predecir y para clasificar. Se describe como un conjunto de tuplas¹⁰ y cada una de ellas se representa como un vector con valores del atributo, cada atributo tiene asociado un conjunto de valores. Los atributos pueden ser nominales, categóricos, o numéricos (límites superior e inferior). La instancia se define como un producto cartesiano de todos los dominios de atributos de entrada. En clasificación, el dominio del atributo objetivo es finito y categórico. Es decir, hay un número finito de clases o categorías para predecir una muestra y son conocidos por el algoritmo de aprendizaje. Un clasificador debe ser capaz de asignar clases a nuevos ejemplos, incluso cuando estos no se han utilizado en el proceso de entrenamiento. La naturaleza de la clasificación es discriminar unos ejemplos de otros, logrando una predicción fiable de la categoría correcta de nuevas instancias. Sin embargo, cuando el atributo de destino está formado por valores infinitos, como en el caso de predecir un número real entre un intervalo, se debe utilizar un modelo de regresión. En síntesis, el objetivo del aprendizaje supervisado es obtener un mapeado de entrada y de salida cuyos valores correctos y definidos son proporcionados por un supervisor.

Seguidamente, se describirá brevemente en que consisten las técnicas de aprendizaje no supervisado siguiendo a García et al. (2015). En este tipo de técnicas no existe supervisor, solo se dispone de datos de entrada. El objetivo, en este caso, es el de encontrar regularidades, irregularidades, relaciones, similitudes y asociaciones en la entrada. En el aprendizaje no supervisado, es posible aprender más y aplicar modelos más complejos que con aprendizaje supervisado. La razón es que en este último se trata de encontrar la conexión entre dos conjuntos de observaciones. En aprendizaje no supervisado además de las técnicas de *clustering* se pueden incluir las siguientes (García et al., 2015):

- *Pattern Mining*: la minería de patrones incluye patrones extraños o negativos. Las reglas de búsqueda de patrones pueden incluir multinivel, multidimensional,

¹⁰ Una tupla es una lista ordenada de elementos.

patrones aproximados, inciertos, comprimidos, raros, negativos y de alta dimensión. Tanto en correlación como en excepción.

- *Outlier Detection*: es el proceso de encontrar ejemplos de datos con comportamientos que son muy diferentes de las expectativas. Estos tienen una alta relación con el análisis de agrupamiento, porque este último encuentra los patrones de la mayoría en un conjunto de datos y organiza los datos en consecuencia, mientras que la detección de valores atípicos intenta detectar aquellos casos excepcionales que presentan desviaciones de los patrones de la mayoría.
- *Técnicas de Process mining (ProM)*: es una técnica de minería de datos que se aplica sobre los registros. Dicha técnica trabaja con un flujo de procesos a partir de los registros realizados por muchas aplicaciones, un ejemplo de aplicación diferencial entre el procedimiento tradicional y la técnica de *Process mining* se puede comprobar en la Figura 27.

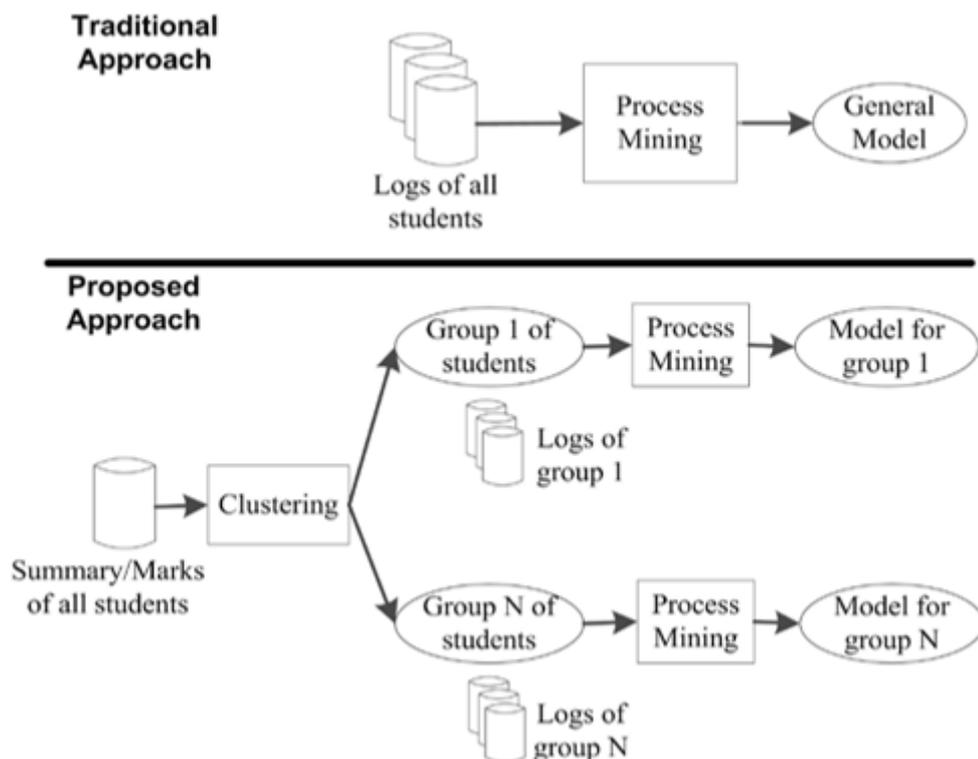


Figura 27

Esquema de análisis de datos con un procedimiento tradicional y con uno de Process Mining. Tomado de Bogarín, Romero, y Cerezo (2016) p. 77. Reproducido con permiso de los autores

Previamente al análisis de los datos, se deberá de realizar una depuración de la base de datos con el fin de eliminar aquellos datos que no son relevantes para el estudio (ver la información recogida en este punto). Seguidamente, se podrá aplicar un algoritmo de ProM y también técnicas de ajuste como *Goodness-of-fit indices* o Índice de ajuste. Estas técnicas indican la diferencia entre el comportamiento observado en el registro y el comportamiento descrito por el modelo de proceso, una secuencia de actividades que pertenecen a un mismo caso se denomina traza, las trazas pueden estar asociadas a rutas de ejecución especificadas por el modelo de minería de procesos. También, podría utilizar EMD de aprendizaje supervisado de clasificación y de regresión. Para ello se pueden emplear distintas herramientas (Bogarín, Romero, y Cerezo, 2017a, 2017b):

- SPSS en sus paquetes de *clustering*.
- WEKA aplicando la técnica de *clustering*.
- ProM (Van Der Aalst, 2011) *Heuristics Miner*, analiza la probabilidad desde el cálculo de frecuencias o relaciones entre las tareas y los constructos de dependencia/frecuencia en tablas y en gráficos.
- *Medidas de la teoría de grafos* (número total de nodos y número total de enlaces) para analizar el nivel de complejidad de los modelos obtenidos.
 - *Intention Mining (IM)* es un análisis de procesos que se focaliza en el análisis del razonamiento a través del análisis de las actividades.
 - *Alpha Miner (AM)* descubre cuál será el mejor algoritmo, si bien las limitaciones de esta técnica es que no usa frecuencias y por ello sólo es adecuado cuando analiza eventos con ruido, es muy poco frecuente en el aprendizaje de datos.
 - *Sequence Pattern Mining (SPM)* es una técnica común en DM y descubre las subsecuencias comunes, encuentra relaciones entre acontecimientos sucesivos. Puede analizar *episode mining (EP)* se basan en los *t-pattern analysis* de los modelos de Markov. SPM se utiliza para analizar las conductas de aprendizaje de los estudiantes.
 - *Graph Mining (GM)* también denominada sub-graph mining hay que diferenciar esta técnica de la de *network analysis (SNA)* esta última puede ser considerada una parte de GM.

- *Genetic algorithm* proporciona modelos de proceso construido sobre matrices causales (entrada y salida) y dependencias para cada actividad. Este enfoque aborda problemas tales como ruido, datos incompletos, constructos de libre elección, actividades ocultas, concurrencia y actividades duplicadas. Los algoritmos genéticos son un método de optimización que maximiza o minimiza una función $f(x_1, x_2, x_3, \dots)$ ->.
- *Fuzzy miner* es un proceso de descubrimiento de algoritmos se utiliza para abordar problemas con un número grande de números y actividades que no están muy estructuradas.
- *Social Network Analysis Technique (SNA)* es una técnica sociométrica que analiza las redes de interacción social, consiste en nodos que representan la organización en entidades y arcos.
- *Visualización de procesos* permite desarrollar una interfaz de visualización del análisis de datos.

Esta forma de análisis permite el seguimiento de cada uno de los grupos detectados y por ende la puesta en marcha de orientaciones personalizadas para cada uno de ellos. Asimismo, la combinación de las técnicas de EDM y de ProM va a servir al docente para estudiar el proceso de comportamiento de todos los estudiantes desde el inicio del desarrollo de la docencia, con el objetivo último de poder adaptarla a las necesidades de cada grupo.

1.9. Técnicas de Machine Learning utilizadas

En los artículos que se presentan en esta tesis se emplearon técnicas de aprendizaje supervisado de clasificación [Máquina de Soporte de Vectores (*Support Vector Machine*), Análisis Discriminante] y de Regresión [Regresión Lineal, Regresión logística, Modelado Lineal Automático, Árboles de Decisión, Redes Neuronales (*neural network*), en concreto de base radial], y de aprendizaje no supervisado de clustering específicamente, *k*-means (ver Figura 26).

1.9.1. Técnicas de aprendizaje supervisado de clasificación

Seguidamente, se presentan algunas de las técnicas de aprendizaje supervisado usadas en los artículos desarrollados en esta tesis.

1.9.1.1. Máquina de Soporte de Vectores (Support Vector Machine)

Esta técnica aplica un conjunto algoritmos que se basan en la teoría del aprendizaje estadístico de Vapnik (2013). En el caso más simple, con sólo dos clases linealmente separables, esta técnica pretende encontrar «la mejor» frontera de separación entre las clases. Entendiendo que es aquella que está más alejada de las dos clases que se quieren separar. Es decir, aquella que «maximiza el margen». Las instancias de cada clase más cercanas a la frontera son los «vectores soporte». La búsqueda de la frontera se resuelve como un problema de optimización en el que se encuentra tanto el hiperplano de separación, como los vectores soporte. Si las clases no son separables, se introduce en el problema de optimización un factor de penalización para minimizar el número de instancias mal clasificadas. Si la separación no es lineal, se puede utilizar una función no lineal que transforma el conjunto de datos de forma que la separación lineal del conjunto transformado se corresponde con una separación no lineal del conjunto de partida. La función de transformación no tiene porque existir, dado que utilizando el «*kernel trick*» es suficiente con tener un producto vectorial en el espacio transformado (Hofmann, Schölkopf, y Smola, 2008), que puede definirse a partir de las instancias de partida, sin que sea necesario conocer de forma explícita la transformación concreta que daría lugar al espacio transformado en el que se aplica el producto vectorial (es decir, basta con obtener el resultado del producto vectorial, sin que sea necesario obtener explícitamente los vectores transformados). Con estas dos extensiones se puede utilizar el método para hacer clasificación, aunque las clases no sean linealmente separables. El método se ha extendido también a problemas multi-clase y a problemas de regresión.

1.9.1.2. Técnica de Análisis Discriminante

Es una técnica de análisis multivariante cuyo objeto es describir si existen diferencias significativas en x grupos de objetos sobre los que se observan y variables discriminantes. Se comparan y se describen las medias de las y variables clasificadoras en los x grupos. Las diferencias halladas intentarán una asignación sistemática en nuevas observaciones y una asignación de estas a grupos ya analizados. Para ello se utilizarán los valores en las y variables que son las clasificadoras. Este procedimiento sería un modelo de predicción de una variable de respuesta categórica a partir de y variables explicativas que normalmente son variables continuas o variables clasificadoras.

1.9.1.3. Técnicas del vecino más cercano (Nearest Neighbor k-nn)

Es un método aprendizaje supervisado que sirve para estimar una función de densidad $F(x/C_i)$ de las variables predictoras x por cada clase C_i . Es un método de clasificación y de regresión (en clasificación devuelve la moda de la clase de los k ejemplos más cercanos y en regresión la media) no paramétrico que estima la función de la densidad de probabilidad de que un elemento x pertenezca a la clase C_i . Este es un método de clasificación de elementos basado en un entrenamiento mediante ejemplos cercanos en el espacio de elementos posibles. La función se aproxima localmente, y por esto se ha denominado a esta técnica, técnica de aprendizaje vago o *lazy learning* (ver Figura 28).

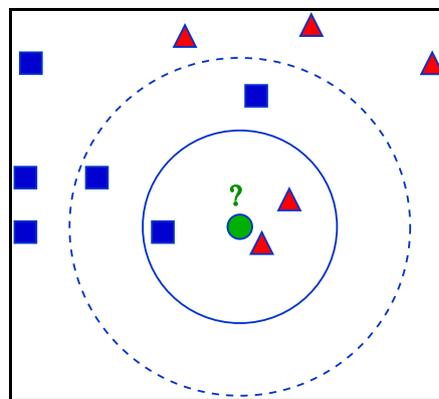


Figura 28

Tomada de Antti Ajanki AnAj

El algoritmo aplicado es el entrenamiento de un conjunto de vectores en un espacio multidimensional, cada ejemplo tiene p atributos desde q clases para la clasificación.

$$x_i = (x_{1i}, x_{2i}, \dots, x_{pi}) \in X$$

Puede utilizar múltiples medidas de distancia, no obstante la más frecuente es la distancia euclidiana en la que un punto es asignado a una clase C si ésta es la más frecuente entre los k ejemplos de entrenamiento más cercano.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^p (x_{ri} - x_{rj})^2}$$

El modelo supone que los ejemplos más cercanos tienen los mismos atributos para la clasificación de un nuevo elemento, aunque es posible que existan atributos ocultos no

contabilizados. Para corregir el sesgo se puede asignar un peso a las distancias de cada atributo, dando más importancia a los atributos considerados más relevantes. También, se puede intentar ajustar los pesos con ejemplos conocidos de entrenamiento, y eliminar los atributos que se consideren irrelevantes.

En el paso de clasificación se aplica la siguiente fórmula

$$\langle x, f((x)) \rangle x \in X$$

donde (x_i) son los elementos que tienen que ser clasificados, los k vecinos más cercanos a i ,

$$\hat{f}(x) \leftarrow \operatorname{argmax}_{v \in V} \sum_{i=1}^k \sigma(v, f(x_i))$$

$$\sigma(a, b) = 1 \text{ si } a = b \text{ y } \sigma(a, b) = 0 \text{ en cualquier otro caso}$$

El valor $\hat{f}(x_q)$ devuelto por el algoritmo es un estimador de $f(x_q)$ entre los k vecinos más cercanos a x_q . La elección de los k depende del n de los datos, los valores altos de k reducen el efecto de ruido en la clasificación, si bien crean límites de clases semejantes. Los mejores valores de k se pueden obtener desde la optimización de uso (Fix y Hodges, 1989). Un ejemplo de procesamiento del vecino más cercano con el paquete estadístico SPSS v.24 se puede observar en la Figura 29. Se analiza la variable tipo de *Blended Learning* empleado en tres opciones de entrenamiento (1 = *Supplemental Blended F2F*; 2 = *Replacement Blended 1*, 3 = *Replacement Blended 2*) sobre las variables de predicción que son en este caso los resultados de aprendizaje en distintas pruebas (Defensa del ABP, Elaboración del ABP y Pruebas tipo test) y en los resultados finales.

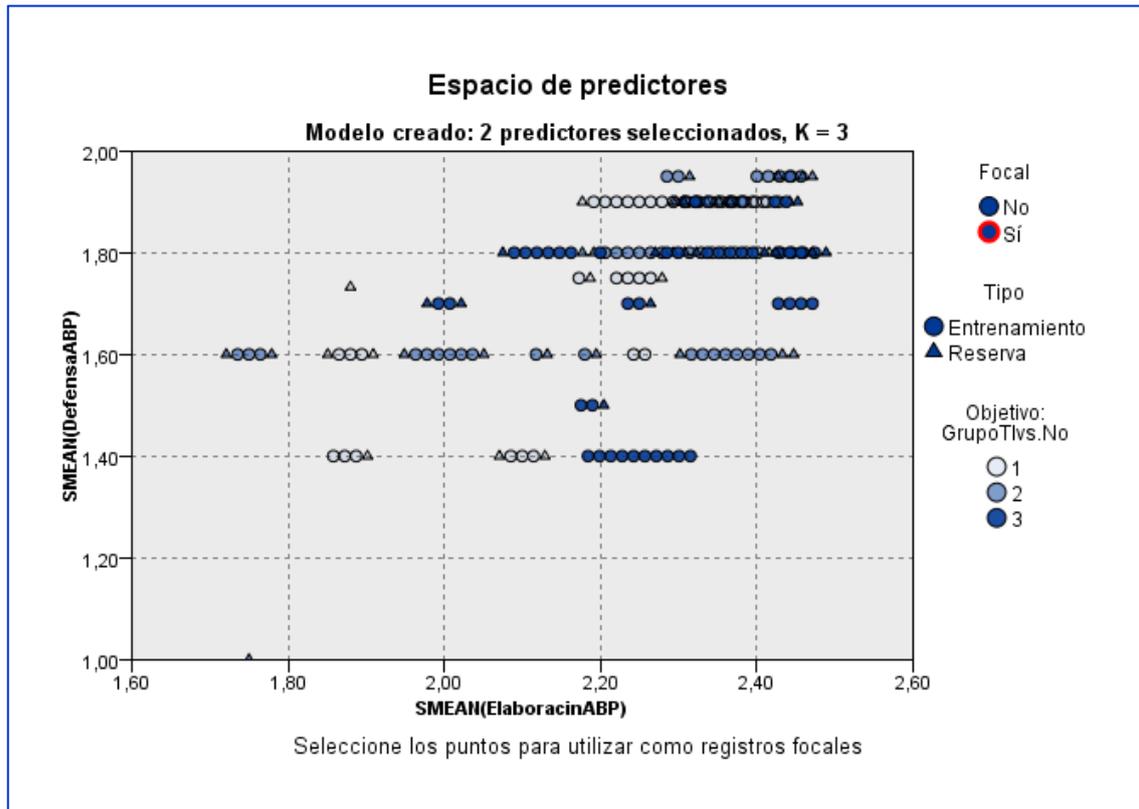


Figura 29

Aplicación de la técnica del vecino más cercano con el paquete estadístico SPSS v.24

El identificador focal permite marcar los casos que tienen un especial interés. Este método utiliza especialmente en casos clínicos, los casos focales muestran los vecinos más próximos a la variable especificada, estos valores siempre son positivos.

1.9.2. Técnicas de aprendizaje supervisado de Regresión

1.9.2.1. Regresión lineal

Es un modelo matemático que se utiliza para aproximar la relación entre dos o más variables, una que se considera la dependiente y otras que se consideran independientes o regresores y una variable de error.

$$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \xi$$

Y_t sería la variable dependiente, las $\beta_i: 1 \leq i \leq n$ serían las pendientes de la recta sobre las variables independientes, n sería el número de ellas, β_0 un desplazamiento y ξ sería el error asociado.

Las rectas de regresión más satisfactorias son las que mejor se ajustan a la nube de puntos o diagrama de dispersión que se genera por la distribución binomial. Para dos variables, la recta de regresión de y sobre x sería:

$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x})$$

Y la de x sobre y

$$x = \bar{x} + \frac{\sigma_{xy}}{\sigma_y^2} (y - \bar{y})$$

Puede hallarse una Regresión Lineal simple, una sola variable dependiente o múltiple, varias variables dependientes.

La correlación r entre las rectas determina el grado de ajuste sería 1 y el peor 0.

Un ejemplo con cuatro variables independientes se puede observar en la Figura 30 y la regresión sobre cada una de las variables independientes se puede comprobar en la Figura 31.

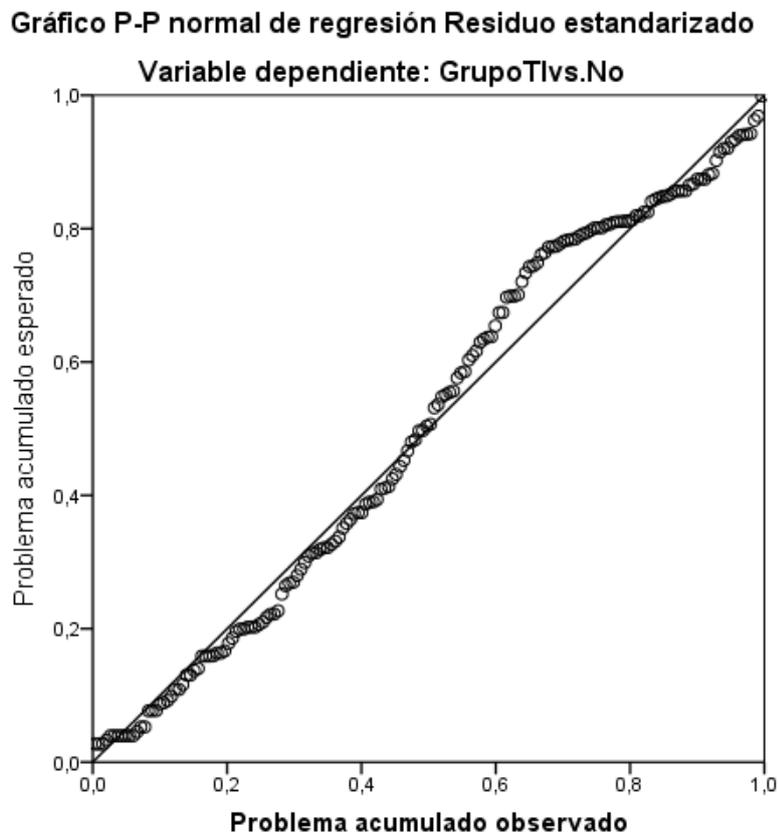


Figura 30

Gráfico de regresión de residuo estandarizado realizado con el paquete estadístico SPSS v.24

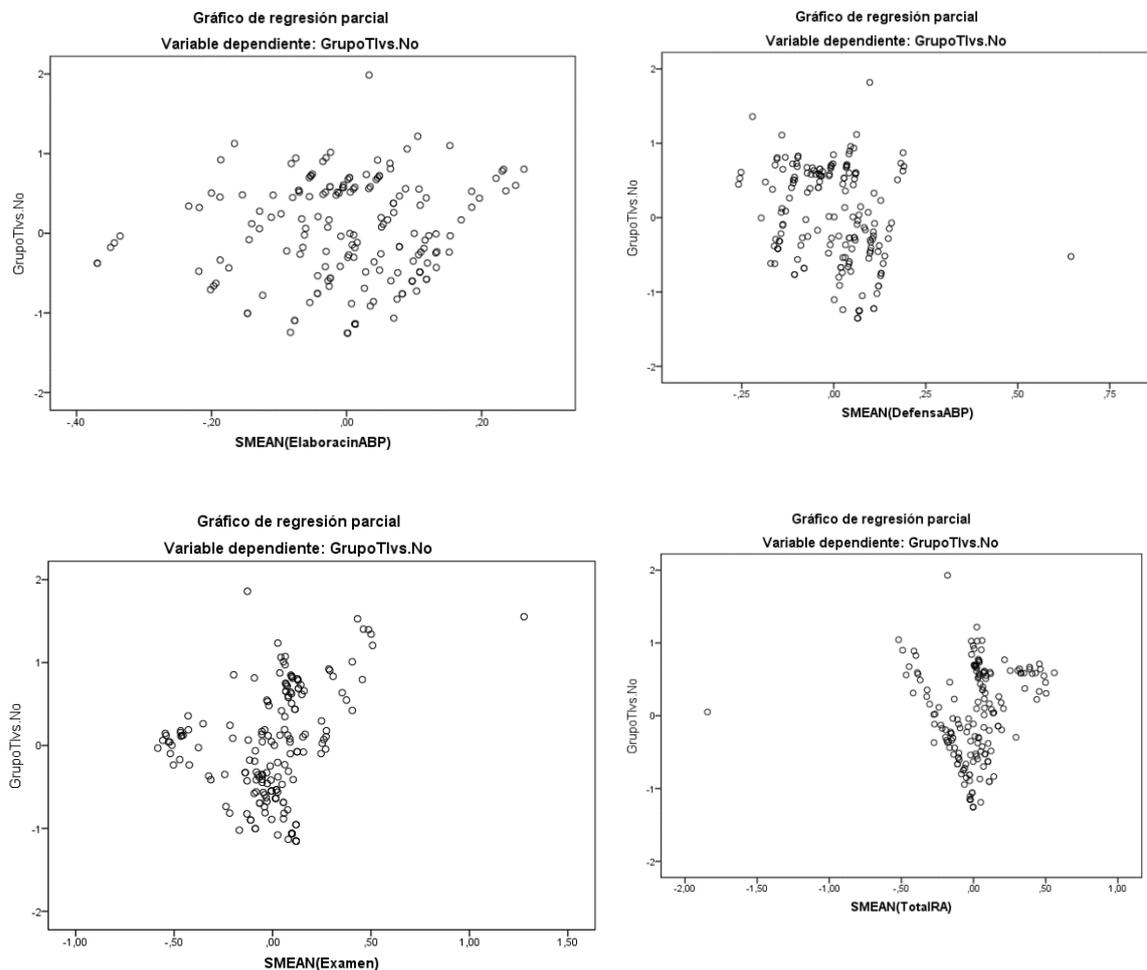


Figura 31

Gráficos de regresión parcial realizado con el paquete estadístico SPSS v.24

Asimismo, hay que tener en cuenta otros dos índices la *Tolerancia*, o probabilidad de entrada, indica si los valores de una variable independiente son redundantes respecto de las otras variables dependientes. Si el valor de alguna de ellas es próximo a 0, habrá que eliminarla ya que dicha variable será una combinación lineal de las otras; y el Valor de Inflación de la Varianza (VIF), cuantifica la multicolinealidad en un análisis de regresión normal de mínimos cuadrados, indica hasta que punto el cuadrado de la desviación estándar estimada se incrementa por razones de colinealidad los valores entre 1-10 se consideran adecuados. Un ejemplo de análisis de Regresión Lineal múltiple realizado con el paquete estadístico SPSS v.24 se presenta en la Figura 32.

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	p	Correlaciones			Estadísticas de colinealidad		
	B	Error estándar				Beta	Orden cero	Parcial	Parte	Tolerancia	VIF
1	(Constante)	-.611	.799		-.765	.446					
	SMEAN(ElaboracinABP)	.524	.420	.113	1.246	.214	.148	.095	.077	.455	2.199
	SMEAN(DefensaABP)	-1.538	.437	-.335	-3.514	.001	-.125	-.260	-.216	.415	2.407
	SMEAN(Examen)	.850	.202	.438	4.204	.000	.550	.306	.258	.347	2.881
	SMEAN(TotalRA)	.226	.201	.169	1.126	.262	.329	.086	.069	.168	5.970

a. Variable dependiente: GrupoSTI

Figura 32

Modelo de Regresión lineal Múltiple realizado con el paquete estadístico SPSS v.24

Nota. Coeficientes no estandarizados = son los coeficientes de regresión parcial que definen la ecuación de regresión en puntuaciones directas. Coeficiente estandarizado = β ayudan a determinar cuál de las variables independientes tiene mayor efecto sobre la variable dependiente y definen la ecuación de regresión en puntuaciones típicas. Colinealidad es la propiedad según la cual un conjunto de puntos están situados sobre la misma línea recta.

1.9.2.2. Regresión logística

Es un modelo de clasificación muy similar a la regresión lineal, solo que en este caso el valor estimado es la probabilidad de pertenencia a una instancia a una clase. Dicho método se utiliza para predecir el resultado de una variable categórica respecto de las variables independientes o variables predictoras. Utiliza modelos lineales generalizados y utiliza una función *logit*. La regresión logística estudia los datos distribuidos de forma binomial

$$Y_i \sim B(p_i, n_j) \text{ donde } i = 1, \dots, m$$

Los n_j son conocidos y las probabilidades de éxito son desconocidas p_i

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \beta_2 X_2 + \dots + \beta_k X_{nk,i}$$

Los parámetros no conocidos de β_j frecuentemente se estiman con el método de Máxima Verisimilitud.

El modelo se explica por:

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1,i} + \beta_2 X_2 + \dots + \beta_k X_{k,i})}}$$

Esta fórmula se puede interpretar como un perceptrón de una capa simple o red neuronal artificial de una sola capa y se calcula con la fórmula:

$$y = \frac{1}{1 + e^{-f(x)}}$$

X es una función analítica que tiene una derivada continua que permite la propagación hacia atrás:

$$y' = y(1 - y) \frac{df}{dx}$$

La regresión logística puede utilizarse para correlacionar la probabilidad de una variable cualitativa binaria que, por ejemplo, puede tomar los valores «1», no ocurrencia de un suceso, y «2», ocurrencia de un suceso, con una variable escalar x que puede explicar el suceso. En este ejemplo, se ha utilizado la variable sexo, al sexo femenino se le ha dado el valor 1 y al sexo masculino el valor 2, se ha hallado una tabla de contingencia entre los valores observados y los esperados, ver Figura 33.

		Sexo = 1		Sexo = 2		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	15	14.489	3	3.511	18
	2	12	14.714	7	4.286	19
	3	15	14.320	4	4.680	19
	4	15	13.141	3	4.859	18
	5	14	12.515	4	5.485	18
	6	13	12.448	6	6.552	19
	7	8	10.914	10	7.086	18
	8	10	9.899	8	8.101	18
	9	8	8.777	11	10.223	19
	10	4	2.782	6	7.218	10

Figura 33

Tabla de contingencia en la prueba de Hosmer y Lemeshow realizado con el paquete estadístico SPSS v.24

1.9.2.3. Modelado lineal Automático

Los modelos lineales predicen un objetivo continuo, basándose en relaciones lineales entre el objetivo y uno o más predictores. Estos modelos son relativamente simples y proporcionan una fórmula matemática fácil de interpretar para la puntuación. Las propiedades de los mismos modelos se comprenden bien y se pueden crear rápidamente en comparación con el resto de los tipos de modelos (como redes neuronales o árboles de

decisión) en el mismo conjunto de datos. Tiene que haber un objetivo y al menos una entrada. El objetivo tiene que ser una variable continua (escala) y no hay ninguna restricción de nivel de medición en los predictores (entradas), los campos categóricos (nominal y ordinal) se utilizan como factores en el modelo y los campos continuos se utilizan como covariables. A modo de ejemplo, se presenta en la Figura 34 la ejecución de un modelo cuyo objetivo es predecir el rendimiento académico de estudiantes universitarios de tercero, en la rama de Ciencias de la Salud, respecto de una serie de variables predictoras: tipo de estrategias metacognitivas utilizadas y accesos en la plataforma al *feedback* del docente y el número de visitas por día realizado, y la pertenencia a distintos grupos así como el trabajo colaborativo en grupos pequeños, los conocimientos previos y la satisfacción de los estudiantes con la práctica docente (ver Figura 35). En este caso las variables que más explican la varianza respecto de los resultados académicos son el tipo de titulación, el grupo clase y el pequeño grupo colaborativo grupo. Asimismo, en las figuras 36 y 37 se presenta respectivamente la curva de residuos y la relación entre los valores pronosticados y los observados.

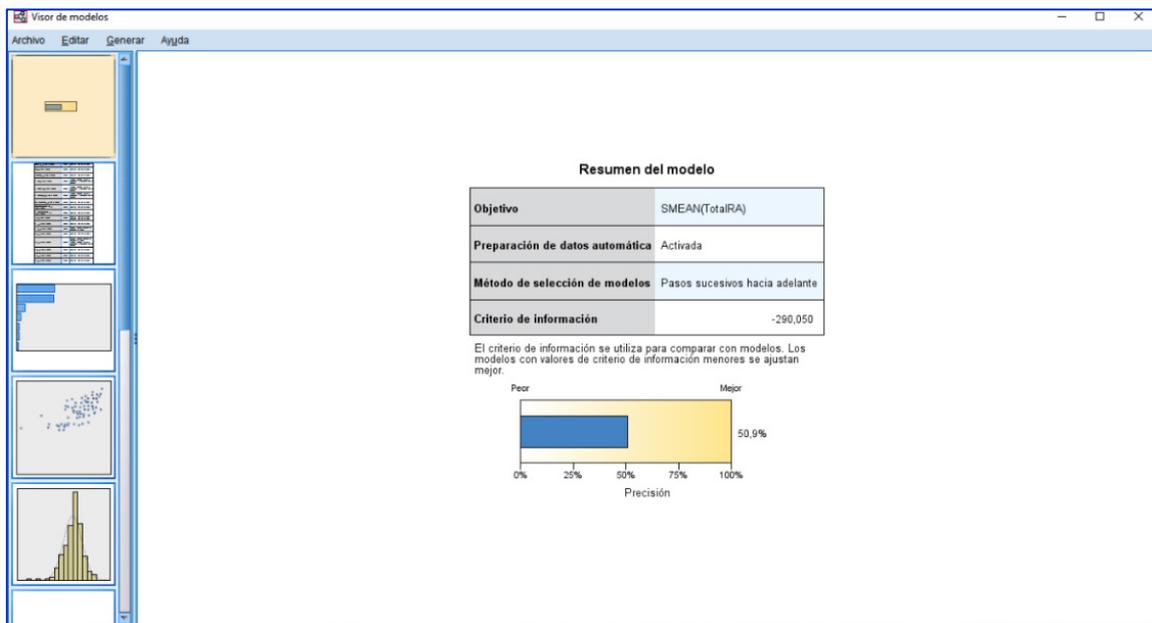


Figura 34

Modelo lineal Automático realizado con SPSS v.24

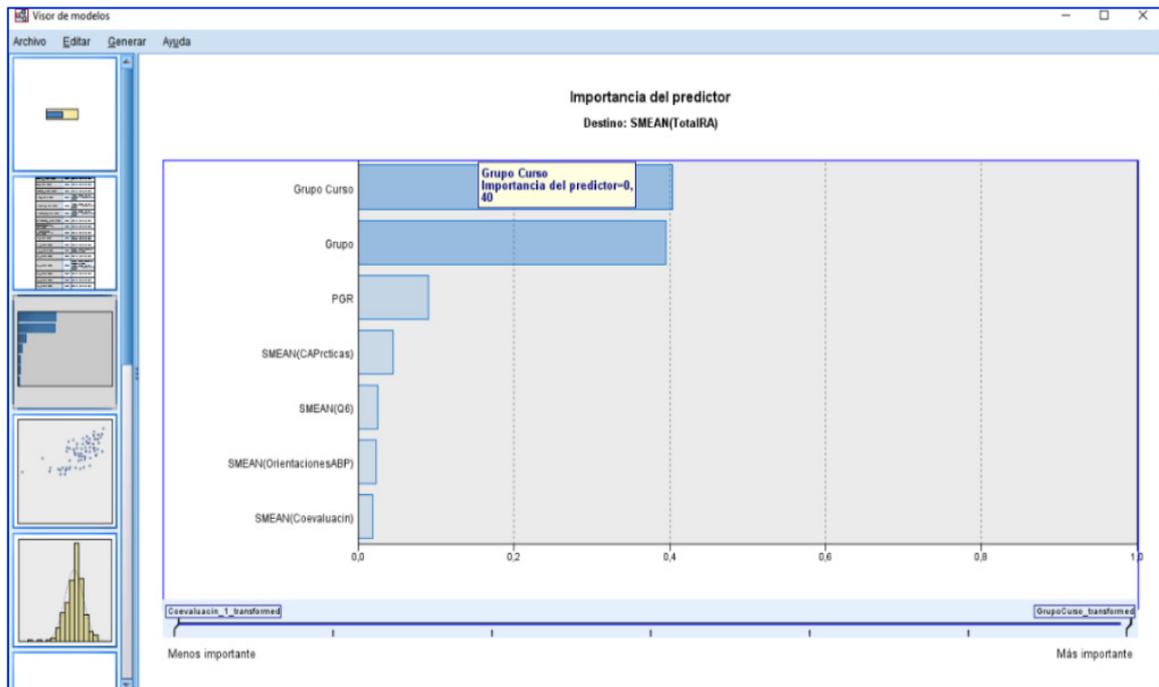


Figura 35

Importancia de los predictores en el modelo lineal automático realizado con SPSS v.24

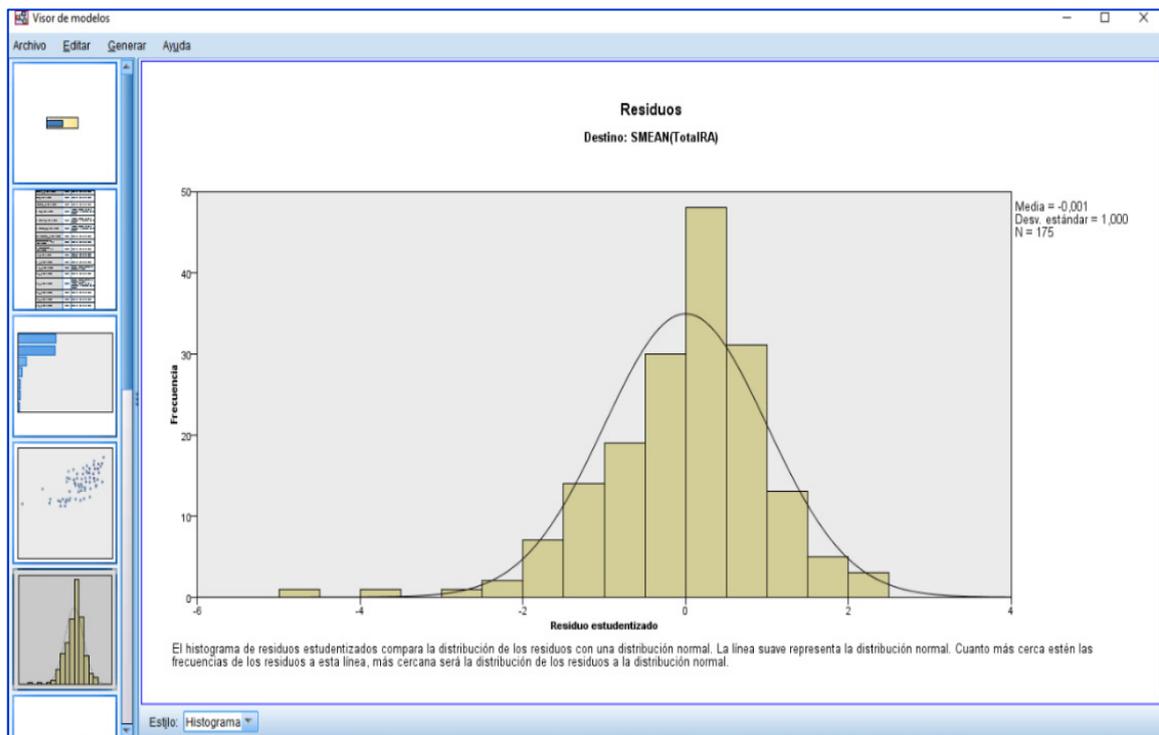


Figura 36

Histograma de residuos en el modelo lineal automático realizado con SPSS v.24

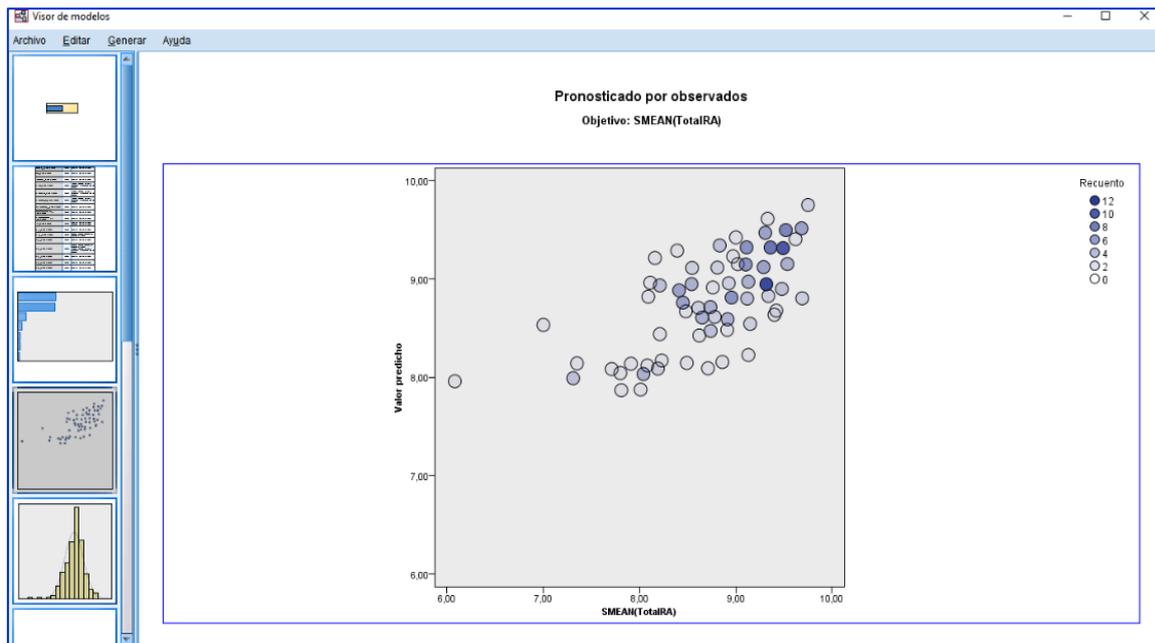


Figura 37

Valores pronosticados por los observados en el Modelo lineal Automático realizado con SPSS v.24

1.9.2.4. Árboles de decisión (*Decision Tree*)

Un árbol de decisión es un algoritmo en el que se utiliza un conjunto de test o preguntas organizadas de forma jerárquica, para guiar el proceso de asignación de clase o calcular el valor de salida. El proceso comienza en el nodo raíz, donde se pide utilizar el valor de uno de los atributos de la instancia a clasificar (o el valor de salida que desea determinar). Dependiendo del resultado de esta comparación, el proceso se dirigirá a una de entre varias ramas (normalmente, sólo a dos), donde se hallará un nodo con una nueva prueba, y nuevas ramas a seguir en función del resultado. El proceso continúa hasta que se alcanza un nodo hoja (un nodo sin más ramas), donde se asigna una clase a la instancia o existe una fórmula para calcular el valor de salida de esa instancia.

El proceso de construcción de un árbol de decisión también comienza en el nodo raíz. En cada nodo se dividirá el conjunto de entrenamiento, por lo que es necesario determinar cuál es el mejor atributo de división y por qué valor se debe hacer dicha división (si se utiliza sólo un valor umbral, la división será de dos subconjuntos, pero sería posible utilizar varios rangos de valores y hacer que el nodo tuviese más de dos ramas). El criterio para la selección de atributos y valores es la optimización de una función (Breiman, 2017; Quinlan, 2014). Una vez seleccionados el atributo y los valores del nodo, se utilizan para determinar a través de qué ramas pasa cada instancia, y en los nodos de llegada correspondientes se repite el

proceso. Dicho proceso concluye cuando el número de instancias es menor que un valor determinado, o cuando se cumple cualquier otro criterio de parada. A los nodos hoja se les asigna la clase mayoritaria de las instancias que han llegado a ese nodo (o se utilizan para calcular una función que dará el valor de salida cuando el árbol se utiliza en problemas de regresión).

Los árboles de decisión son muy populares en la minería de datos y en el aprendizaje automático por varias razones: son rápidos de construir, interpretables e inestables (es decir, pequeños cambios en el conjunto de entrenamiento darán como resultado árboles muy diferentes). Esta última propiedad ha hecho que se utilicen profusamente para la construcción de «multi-clasificadores» (Maudes et al., 2012; Díez-Pastor et al., 2014) y de regresores (Pardo et al., 2013; Arnaiz-González, Díez-Pastor, Rodríguez, y García-Osorio, 2016b). Un ejemplo de aplicación de árboles de decisión realizado con SPSS se puede consultar en la Figura 38 y en la Figura 39.

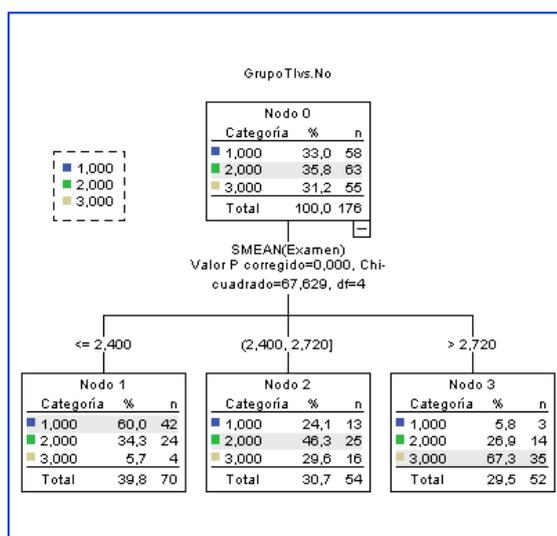


Figura 38

Árbol de decisión realizado con el paquete estadístico SPSS v.24

Nodo	1		2		3		Total		Categoría pronosticada	Nodo padre	Variable	p ²	Chi-cuadrado	gl	Valores de división
	N	Porcentaje	N	Porcentaje	N	Porcentaje	N	Porcentaje							
0	58	33,0%	63	35,8%	55	31,3%	176	100,0%	2						
1	42	60,0%	24	34,3%	4	5,7%	70	39,8%	1	0	SMEAN(Examen)	.000	67,629	4	≤ 2,400
2	13	24,1%	25	46,3%	16	29,6%	54	30,7%	2	0	SMEAN(Examen)	.000	67,629	4	(2,400, 2,720]
3	3	5,8%	14	26,9%	35	67,3%	52	29,5%	3	0	SMEAN(Examen)	.000	67,629	4	> 2,720

Figura 39

Resultados del Árbol de decisión realizado con el paquete estadístico SPSS v.24

1.9.3. Técnicas de aprendizaje no supervisado

Las técnicas de agrupamiento o *clustering* se utilizan cuando no hay información de clase que predecir, pero los datos se tienen que dividir en grupos naturales o clústeres. Dichos grupos reflejan subgrupos de ejemplos que comparten algunas propiedades o tienen algunas similitudes. Trabajan calculando una medida de distancia multivariada entre observaciones que están más estrechamente relacionadas (ver definición aprendizaje no supervisado p. 28). Seguidamente, se van analizar las técnicas de *k*-means y las Redes de función de base radial (RBF).

1.9.3.1. Utilización de la técnica de clustering de *k*-means

Sea un conjunto X y una medida de distancia $d: X \times X \rightarrow \mathbb{R}$. La salida del algoritmo *k*-means es un conjunto de centros $C = \{c_1, c_2, \dots, c_k\}$ que están definiendo implícitamente un conjunto de clústeres en el que cada punto pertenece al clúster representado por el centro más cercano, $\Phi_c(x) = \operatorname{argmin}_c d(x, c)$ y del objetivo es buscar el conjunto C que minimice:

$$\sum_{x \in X} d(\Phi_c(x), x)^2$$

Lo que implica que cada punto se asignará al centro más cercano para lo que se minimizará el cuadrado de las distancias de los puntos al centro asignado.

Asimismo, el método *k*-means++ que tiene una calidad de optimización local. La inicialización inteligente o Smart implica la elección de un primer centro de clúster de forma uniforme desde la aleatorización de los puntos de los datos. Desde cada observación (x) se computa la distancia $d(x)$ hacia el centro del clúster más cercano. Seguidamente se elige un nuevo centro del clúster desde la media de los puntos de los datos con una probabilidad de x empezando a elegir la proporcionalidad de $d(x)^2$ posteriormente se repite los pasos dos y tres desde los k centros elegidos (Fox y Guestrin, 2018). El problema es computacionalmente difícil (NP-hard). Sin embargo, hay eficientes heurísticas que se emplean comúnmente y convergen rápidamente a un óptimo local. Estos suelen ser similares a los algoritmos expectation-maximization de mezclas de distribuciones gaussianas por medio de un enfoque de refinamiento iterativo empleado por ambos algoritmos. Además, los dos algoritmos usan los centros que los grupos utilizan para modelar los datos. Sin embargo, *k*-means tiende a encontrar grupos de extensión espacial comparable, mientras que el mecanismo expectation-maximization permite que los grupos tengan formas diferentes.

Dado un conjunto de observaciones o valores (x_1, x_2, \dots, x_n) cada una de ellas es un vector real de d dimensiones, k -means realiza una partición de las observaciones en k conjuntos ($k \leq n$) para minimizar la suma de los cuadrados dentro de cada grupo $S = \{S_1, S_2, \dots, S_k\}$ donde μ_i es la media de puntos en S_i .

$$\operatorname{argmin}_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2$$

El algoritmo que se utiliza es una técnica de refinamiento iterativo o algoritmo Lloyd. Se parte de un conjunto inicial de k centroides $m_1^{(1)}, \dots, m_k^{(1)}$, el algoritmo se alterna en dos pasos.

Paso de asignación, en el que se asigna cada observación al grupo con la media más cercana, implica la partición de las observaciones respecto del diagrama de Voronoi¹¹, generado por los centroides.

$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\| \leq \|x_p - m_j^{(t)}\| \forall 1 \leq j \leq k\}$$

Cada x_p se sitúa en un centroide $S_i^{(t)}$ aunque podría situarse en dos pasos:

Paso de actualización, se calculan los nuevos centroides como el centroide de las observaciones dentro del grupo. El algoritmo se considera que ha convergido cuando las asignaciones no cambian.

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

En las figuras siguientes se presenta la demostración del algoritmo.

¹¹ Es un método de interpolación basado en la distancia euclidiana que es especialmente adecuada cuando los datos son cualitativos. Se unen los puntos entre las distancias trazando las mediatrices de los segmentos de unión, las intersecciones de las mediatrices determinan unos polígonos en un espacio bidimensional alrededor de un conjunto de puntos de control, el perímetro de los polígonos generados es equidistante a los puntos vecinos en un área de influencia.

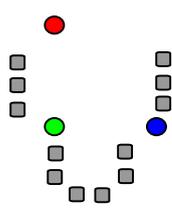


Figura 40
Centroides iniciales en este caso $k = 3$ se generan de forma aleatoria dentro de un conjunto de datos. Tomado de Weston.1 (2018)

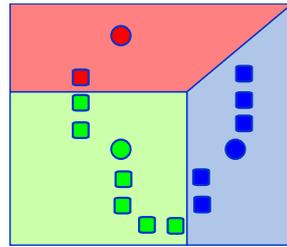


Figura 41
Los k grupos son generados asociando el punto con la media más cercana. Se realiza a través del diagrama de Voronoi generado por los centroides Weston.2 (2018)

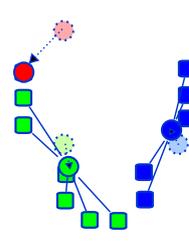


Figura 42
El centroide de cada uno de los k grupos se recalcula Weston.3 (2018)

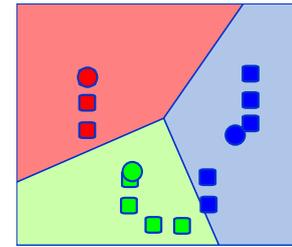


Figura 43
Los pasos 2 y 3 se repiten hasta que se logra la convergencia. Weston.4 (2018)

Las características de k -mean que hacen que este algoritmo sea eficiente también pueden generar problemas. Uno de ellos es la distancia euclidiana, la varianza es una medida de dispersión de los grupos. El número de grupos k es un parámetro de entrada y es necesario una elección adecuada del número de los grupos en un conjunto de datos. El ejemplo de Weston indica la estructura de k grupos, las cruces son los centroides y el algoritmo converge en cinco interacciones (ver Figura 44).

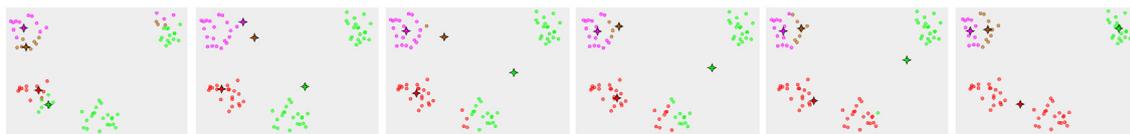


Figura 44

Ejemplo de convergencia del algoritmo k -means en cinco interacciones, tomado de Agor (2018).

Una limitación de k -means es el modelo de agrupamiento, ya que esta técnica tiende a formar grupos de tamaños parecidos y en cambio un algoritmo de expectación-maximización (EM) alterna pasos de esperanza de la verosimilitud (E) mediante la inclusión de variables latentes y un paso de maximización (M) donde se computan estimadores de máxima verosimilitud mediante la maximización de la verosimilitud esperada en el paso E, los parámetros que se encuentran en el paso M se emplean para comenzar el paso E siguiente y así sucesivamente.

En la Figura 45 se puede comprobar las comparaciones de agrupamiento entre un modelo de clustering *k-mean* y uno EM.

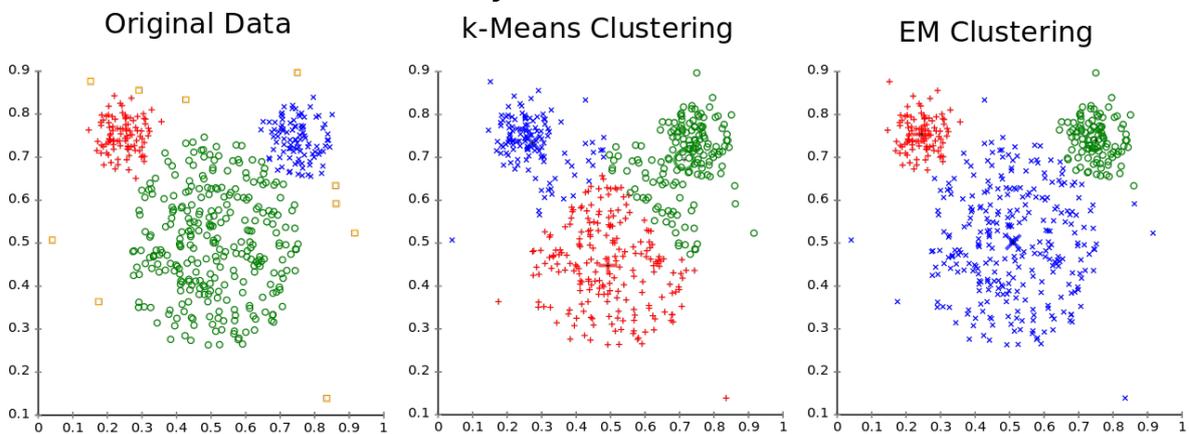


Figura 45

Comparación de agrupamiento entre un modelo de clustering *k-mean* y uno EM, tomado de Chire (2018).

El algoritmo de EM, es un algoritmo iterativo, en un primer paso estima los clústeres posibles respecto de los parámetros presentes estimados.

$$\hat{r} = \frac{\pi N(x_i | \hat{\mu}_k, \hat{\epsilon}_k)}{\sum_{j=1}^k \hat{\pi}_j N(x_i | \hat{\mu}_j, \hat{\epsilon}_j)}$$

Seguidamente, se realiza Maximum Likelihood sobre los parámetros que presentan responsabilidades.

$$\hat{\pi}_k, \hat{\mu}_k, \hat{\Sigma}_k | \{r_{ik}, x_i\}$$

1.9.3.2. Utilización redes neuronales

Las redes neuronales son un modelo computacional que se basa en un conjunto grande de unidades neuronales simples (neuronas artificiales) que pretende imitar el funcionamiento cerebral humano. Cada neurona artificial está conectada con otras muchas y los enlaces entre ellas pueden aumentar o disminuir la conexión con neuronas contiguas. Asimismo, cada unidad neuronal individual opera empleando funciones de suma. Los sistemas aprenden y se forman a ellos mismos. Las redes neuronales tienen varias capas. Las

redes neuronales dinámicas son las más actuales y pueden realizar nuevas conexiones y unidades neuronales. Hay pues, una entrada de información (*input*) un procesamiento, establecimiento de relaciones neuronales en distintas capas, y una salida (*output*). La función de red de una neurona es $f(x)$ que está compuesta de otras funciones $j(x)$, es una estructura de red en las que existen relaciones entre otras variables es una suma ponderada no lineal en la que $f(x) = g(\sum_i w_i j(x))$, g sería la función de activación y es una función predeterminada, como la tangente hiperbólica o la función sigmoide. Esta función proporciona una transición suave entre el valor de entrada y el valor de salida. Se pueden diferenciar las Redes neuronales de perceptrón multicapa y las Redes neuronales de base radial.

a. Redes neuronales de Perceptrón Multicapa

Las Redes Neuronales Artificiales (RNA) de tipo Perceptrón Multicapa (PM) se encuentran entre las arquitecturas de red más poderosas y populares. Están formadas por una capa de entrada, un número arbitrario de capas ocultas y una capa de salida. Cada una de las neuronas ocultas o de salida recibe una entrada de las neuronas de la capa previa (conexiones hacia atrás). Si bien, no existen conexiones laterales entre las neuronas dentro de cada capa. La capa de entrada contiene tantas neuronas como categorías correspondan a las variables independientes que se desean representar. La capa de salida corresponde a la variable respuesta que en este ejemplo es una variable categórica (Longoni, Porcel, López, y Dapozo, 2010). Tiene la capacidad de resolver problemas que no son linealmente separables. La capa de entrada está formada por N neuronas y la oculta está conformada por M , donde M suele ser menor que N , por lo que se tiene que transformar un vector de N componentes en uno de M componentes. El valor óptimo de M es dependiente del problema. Tiene el problema de la comprensión y descomprensión de los vectores, donde cuanto mayor sea el factor de comprensión mayor será el error. Un ejemplo de realización de una RNA de PM se puede comprobar en la Figura 46.

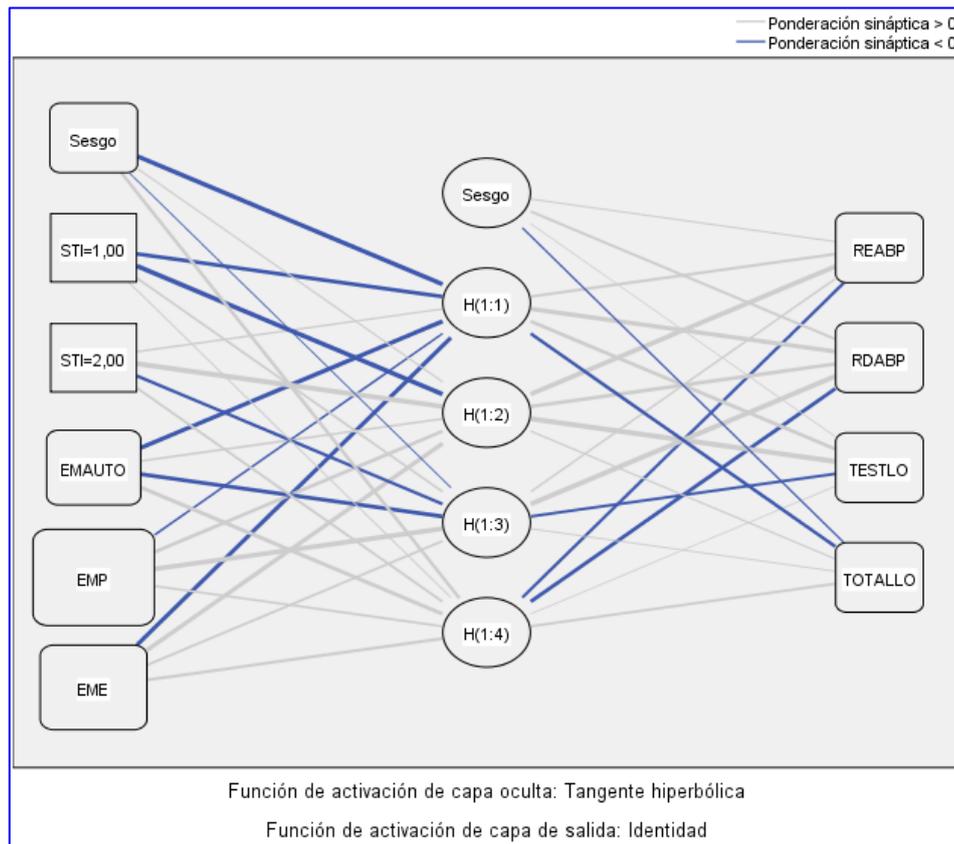


Figura 46

RNA de PM realizado con el paquete estadístico SPSS v.24

Tiene una entrada en la que se contemplan las variables de resultados de aprendizaje de los estudiantes en distintos procedimientos de evaluación: elaboración de un proyecto, defensa de un proyecto, pruebas tipo test y totales, una capa oculta y una capa de salida en la que se contempla el tipo de *Blended Learning* utilizado y las estrategias metacognitivas empleadas por los estudiantes. En concreto, en este caso, las variables en orden de importancia fueron la utilización de estrategias metacognitivas de planificación (ver Figura 47). También permite analizar la dispersión de las puntuaciones pronosticada (ver Figura 48) y su relación con el residuo en cada variable (ver Figura 49).

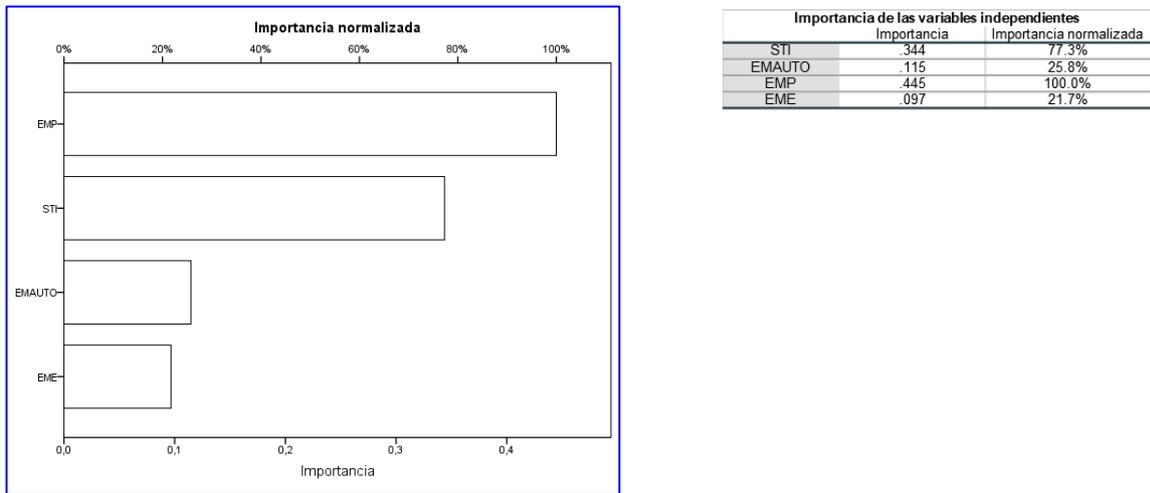


Figura 47

Importancia de las variables en RNA de PM utilizando el paquete estadístico SPSS v.24

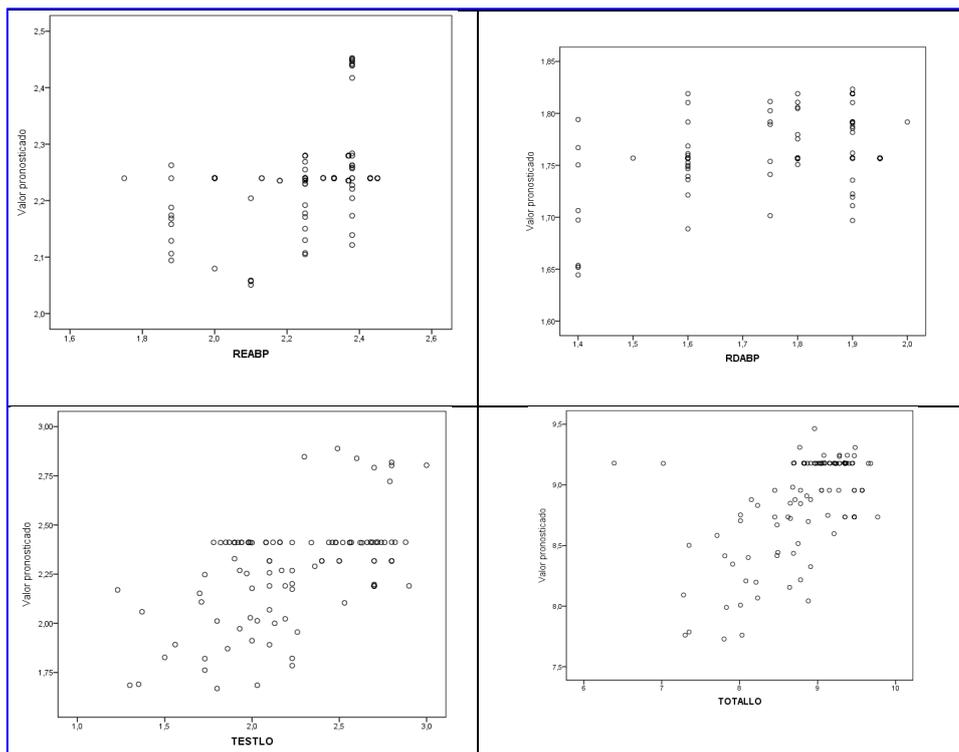


Figura 48

Valores pronosticados de las variables en RNA de PM utilizando el paquete estadístico SPSS v.24

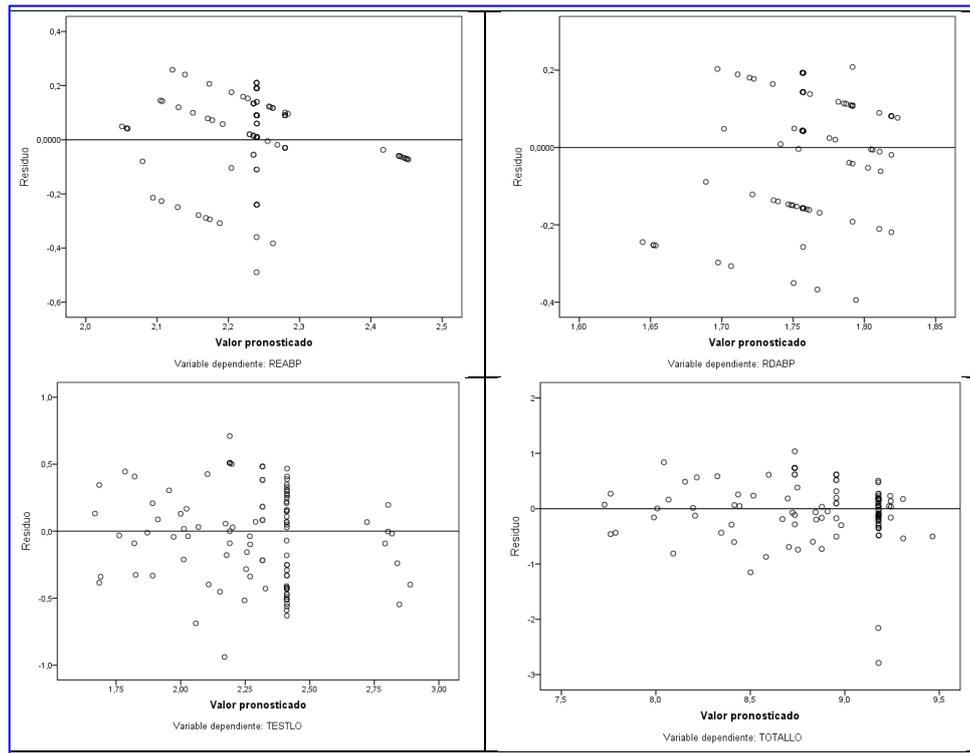


Figura 49

Histograma de residuos en el modelo lineal automático realizado con SPSS v.24

b. Redes neuronales de base radial (RNBR)

Las Redes Neuronales Artificiales (RNA) de Función de Base Radial (FBR) son aquellas cuyas funciones de activación en los nodos ocultos son radialmente simétricas. Se dice que una función es radialmente simétrica o es una Función de Base Radial (FBR), si su salida depende de la distancia entre un vector que almacena los datos de entrada y un vector de pesos sinápticos que recibe el nombre de centro o centroide. Las redes FBR presentan tres capas de conexión hacia adelante: la capa de entrada, la capa oculta o intermedia y la capa de salida. Las neuronas de la capa de entrada, simplemente envían la información a la capa intermedia. Las neuronas de la capa oculta se activan en función de la distancia que separa cada patrón de entrada con respecto al centroide que cada neurona oculta almacena a la que se le aplica una función radial con forma gaussiana. Las neuronas de la capa de salida son lineales y simplemente calculan la suma ponderada de las salidas que proporciona la capa oculta (Longoni et al., 2010). Asimismo, estas calculan la salida de la función en función de la distancia a un punto que se denomina centro, tenemos pues un vector de entrada respecto de un centro.

$$f(x) = (\|x - c_i\|)$$

A cada neurona de la capa de entrada le corresponde una función de base radial ($\varphi(x)$) y un peso de salida w el patrón de salida se dirige a una neurona de salida que suma las entradas y ejecuta una salida.

$$F(x) = \sum_{i=1}^N w_i \varphi(\|x - c_i\|)$$

Es importante señalar que el modelo híbrido de base radial utiliza con más frecuencia el algoritmo *k-mean*, donde el número de clústeres sería el número de neuronas ocultas en la red de base radial.

Un ejemplo realizado con el paquete estadístico SPSS v.24 se puede observar en la Figura 50.

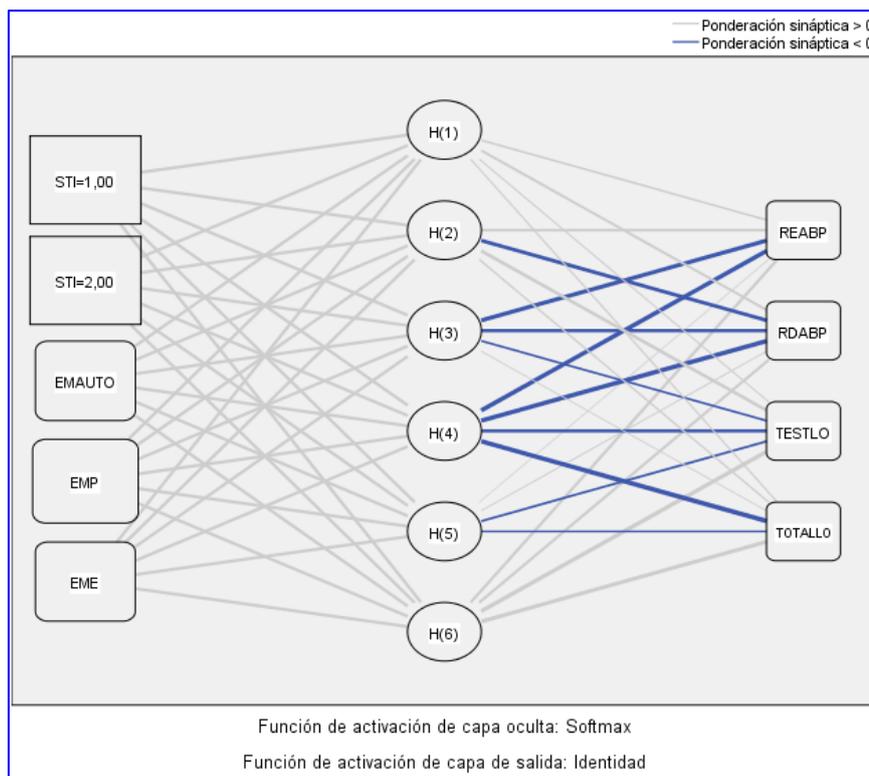


Figura 50

RNA de BR realizado con el paquete estadístico SPSS v.24

En este caso las variables pronosticadas como mejores predictoras de los resultados de aprendizaje fueron el tipo de *Blended Learning* utilizado, las estrategias metacognitivas de planificación, de autoconocimiento y de evaluación (ver Figura 51).

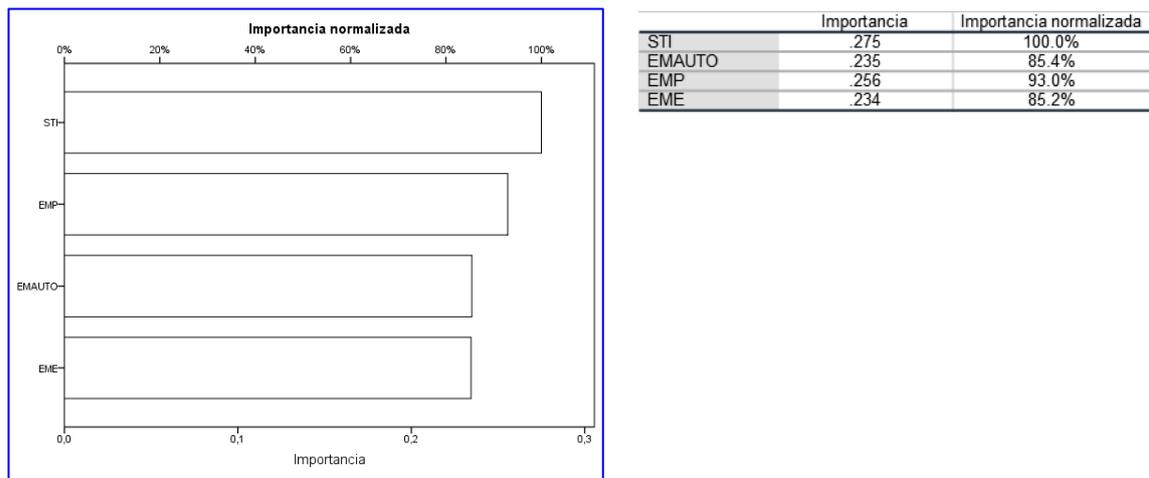


Figura 51

Importancia de las variables en RNA de BR utilizando el paquete estadístico SPSS v.24

En la Figura 52, se puede comprobar los valores pronosticados de las puntuaciones y en la Figura 53 los residuos.

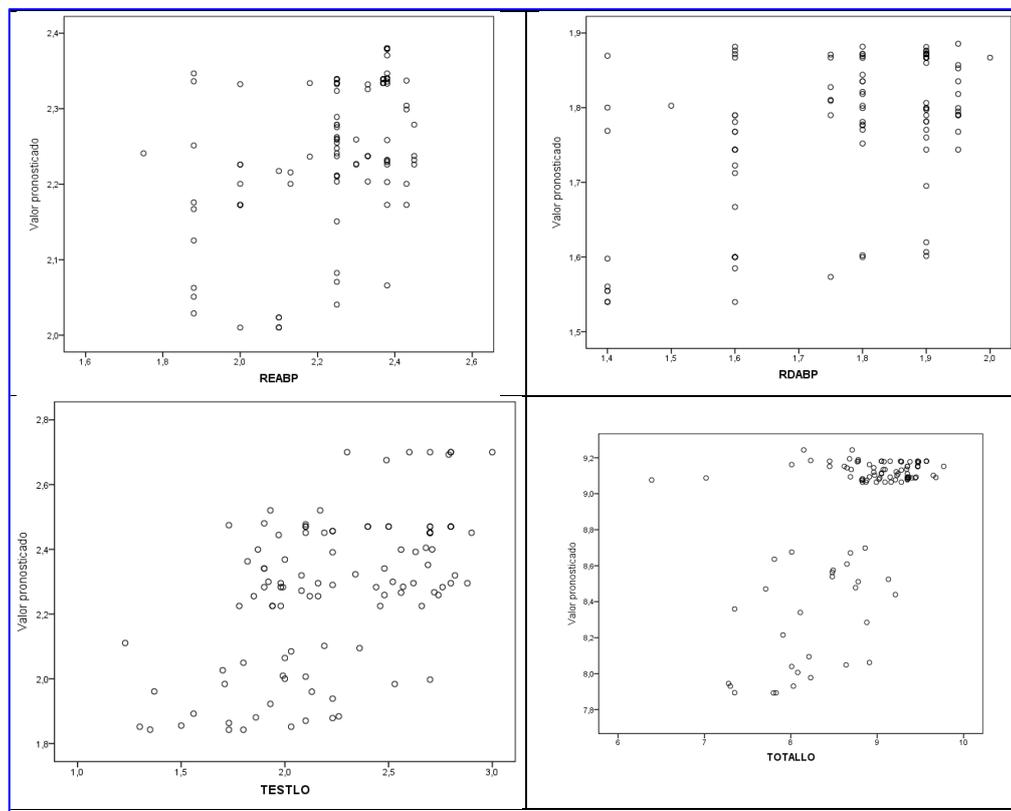


Figura 52

Valores pronosticados de las variables en RNA de BR utilizando el paquete estadístico SPSS v.24

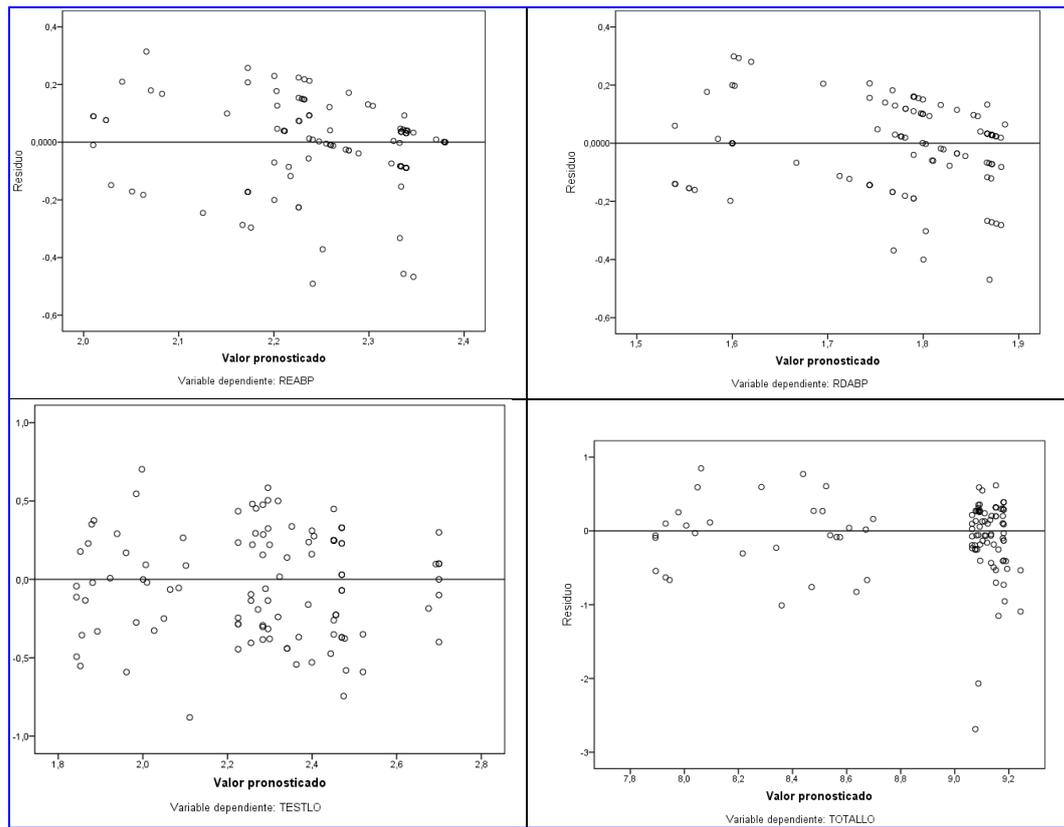


Figura 53

Valores pronosticados y residuos de las variables en RNA de BR utilizando el paquete estadístico SPSS v.24

En síntesis, los análisis de clasificación, regresión y *clustering* que permiten las técnicas de *Machine Learning* son esenciales en la aplicación al campo de estudio de la Ciencia Cognitiva (Hu, Wen, Chua, y Li, 2014; Jones, 2014). En concreto, en el campo del análisis conductual de los patrones de aprendizaje en estudiantes que es el objeto de este trabajo de investigación.

2. PUBLICACIONES Y ACCIONES DE DIFUSIÓN DE LA INVESTIGACIÓN

2.1. Publicaciones principales

Se presentan como méritos principales de esta tesis doctoral cinco publicaciones 4 en la Web of Science dos Q1, dos Q2 (una de ellas en prensa) y una en Scopus (SJR Q3). Se pueden consultar en el Apéndice I.

(a) [Artículo 1](#)

Sáiz, M.C., Marticorena, R., García-Osorio, C.I., & Díez-Pastor, J.F. (2017). How Do B-Learning and Learning Patterns Influence Learning Outcomes? *Frontiers in Psychology*, 8(745), 1-13. doi: 10.3389/fpsyg.2017.00745.

Estándares de calidad de la publicación:

Estándares de calidad de la revista: *Frontiers in Psychology* es una revista de acceso abierto que está indexada en la Web of Science en las categorías de Psychology y Multidisciplinary en SSCI. La revista en la categoría de Psychiatry and Psychology está posicionada en un Q1 (25/774) y en un Q2 la categoría de Multidisciplinary. Asimismo, tiene un factor de impacto de 2.089. Además, tiene un percentil de 71.481 39/135. Datos extraídos a fecha 19 de junio de 2019.

ISSN: 1664-1078 y eISSN: 1664-1078 en Switzerland

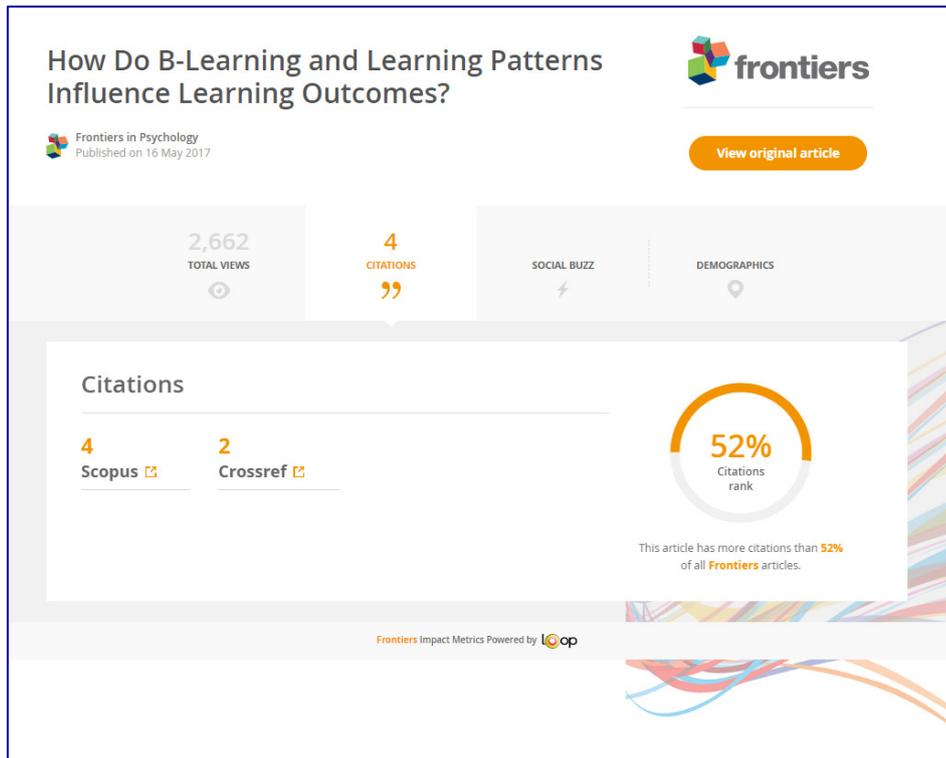
Estándares de calidad del artículo: citas, vistas y descargas:

Citas

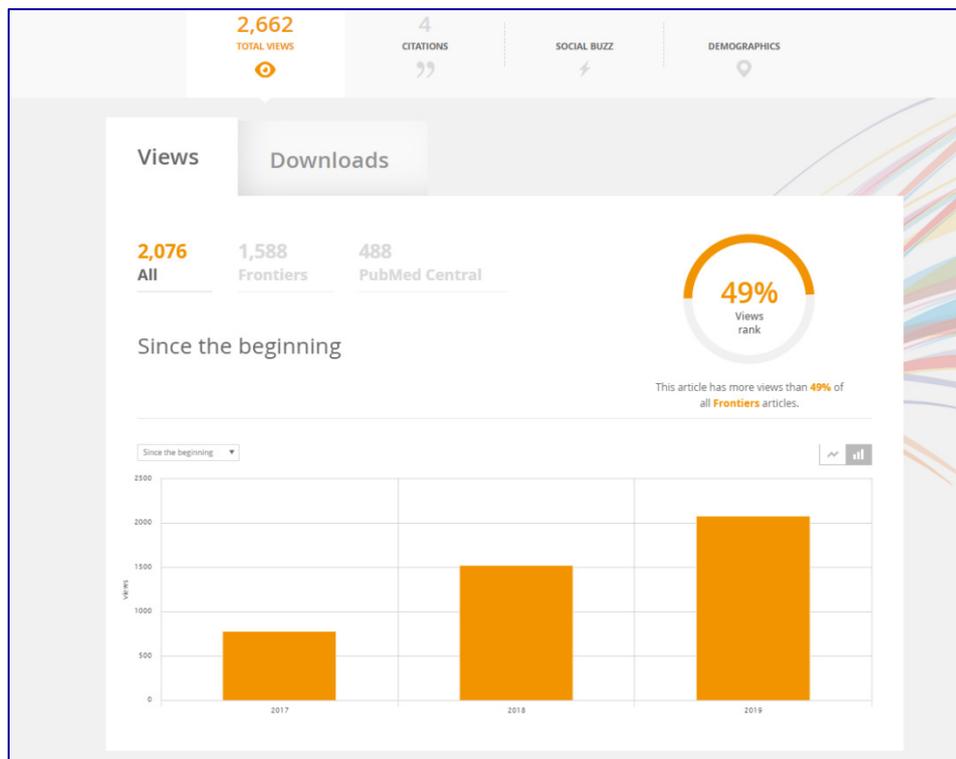
En la web of Science 4, en el loop de *Frontiers in Psychology* 3, en Scopus 3.

Citas en Google scholar 10

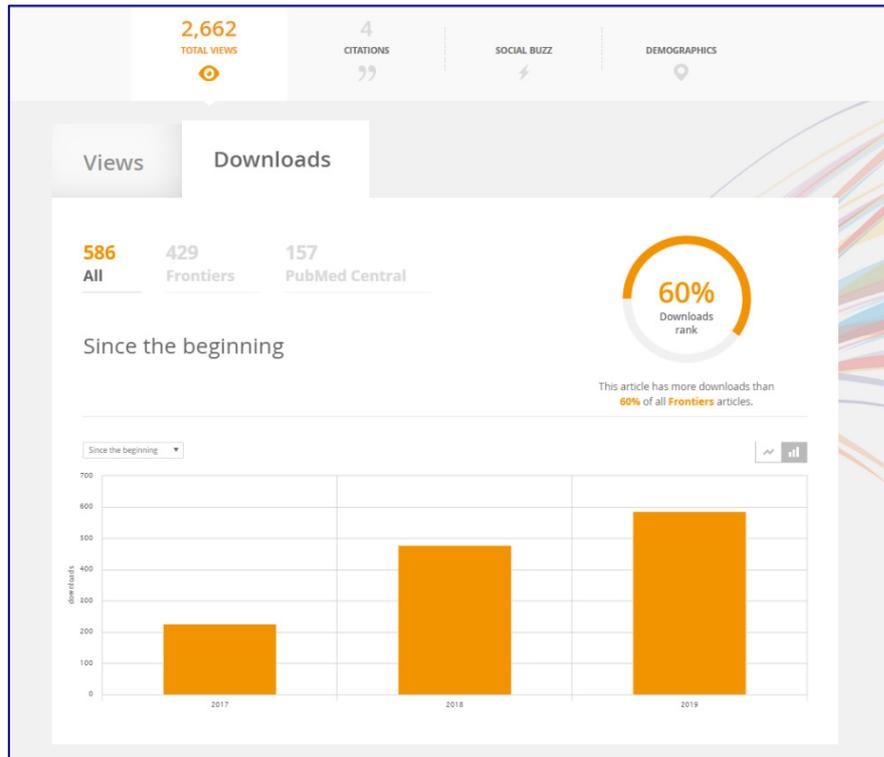
<https://scholar.google.es/scholar?oi=bibs&hl=es&cites=7476101076150529782>



Vistas



Descargas



(b) [Artículo 2](#)

Sáiz, M.C., Marticorena, R., García-Osorio, C.I., & Díez-Pastor, J.F. (2019). Does the Use of Learning Management Systems With Hypermedia Mean Improved Student Learning Outcomes? *Frontiers in Psychology*, 10(88), 1-14. doi: 10.3389/fpsyg.2019.00088. Publicado dentro del topic de *Psychology, Technological Innovation, and Entrepreneurship*.

Estándares de calidad de la revista: Frontiers in Psychology es una revista de acceso abierto que está indexada en la Web of Science en las categorías de Psychology y Multidisciplinary en SSCI. La revista en la categoría de Psychiatry and Psychology está posicionada en un Q1 (25/774) y en un Q2 la categoría de Multidisciplinary. Asimismo, tiene un factor de impacto de 2.089. Además, tiene un percentil de 71.481 39/135. Datos extraídos a fecha 19 de junio de 2019.

ISSN: 1664-1078 y eISSN: 1664-1078 en Switzerland

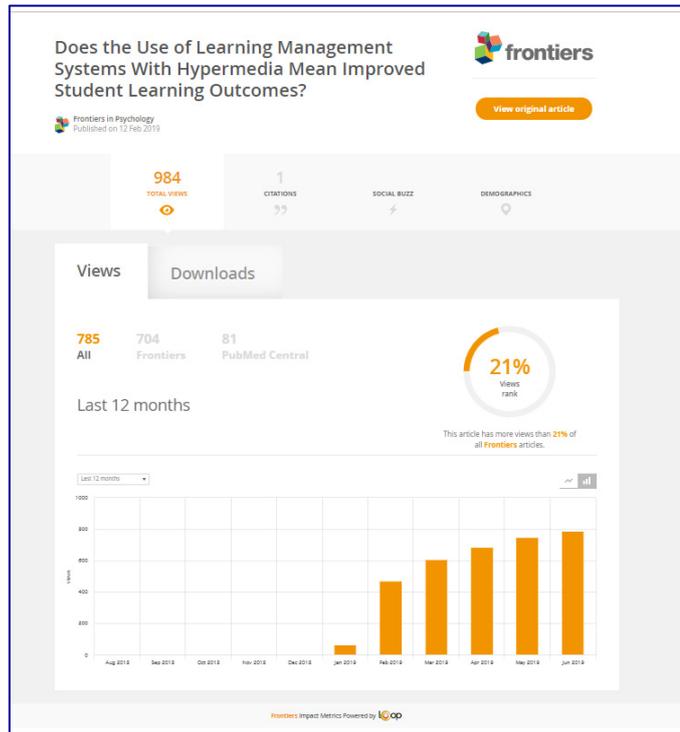
Estándares de calidad del artículo: citas, vistas y descargas:

Citas

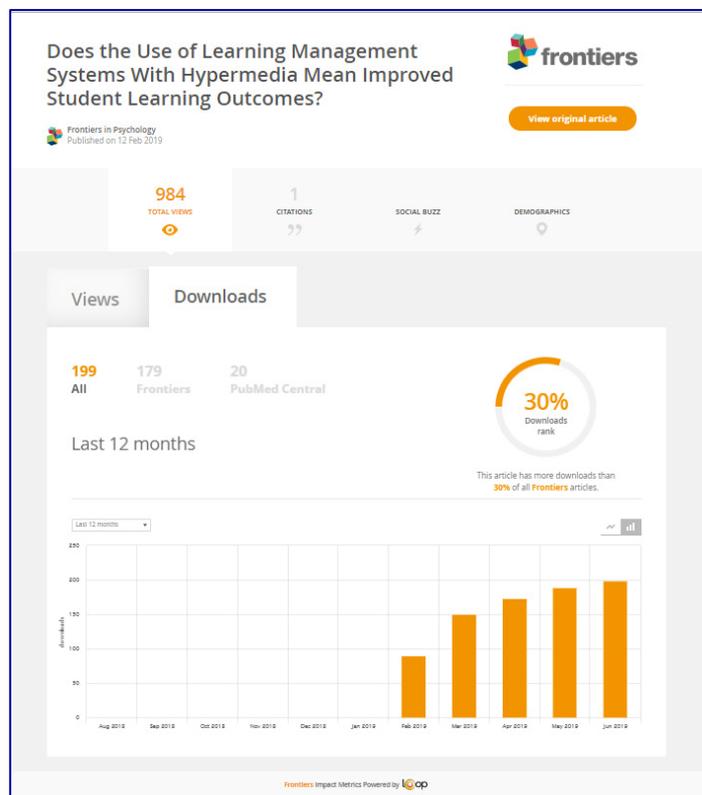
En Scopus 1.



Vistas



Descargas



(c) [Artículo 3](#)

Sáiz, M.C., García-Osorio, C.I., & Díez-Pastor, J.F. (2019). Differential efficiency of the resources used in b-learning environments. *Psicothema*, 31(2), 170-178. doi: 10.7334/psicothema2018.330.

Estándares de calidad de la revista: Psicothema es una revista de acceso abierto que está indexada en la Web of Science en las categorías de Psychology y Multidisciplinary en SSCI. La revista en la categoría Psychology Multidisciplinary está posicionada en un Q2 (285/774) y en un Q2 la categoría de Psychiatry/ Psychology. Asimismo, tiene un factor de impacto de 1.516. Además, tiene un percentil de 58.148. Datos extraídos a fecha 19 de junio de 2019.

ISSN: 0214-9915 y eISSN: 1886-144X en Oviedo (España).

Estándares de calidad del artículo: citas, vistas y descargas: es un artículo de reciente publicación por lo que aún no se han registrado.

(d) [Artículo 4](#)

Sáiz-Manzanares M.C., García Osorio C.I., Díez-Pastor J.F., Martín Antón. (2019). Will personalized e-Learning increase deep learning in higher education? *Information Discovery and Delivery*, 47(1), 53-63. doi: 10.1108/2fidd-08-2018-0039.

Estándares de calidad de la revista: *Information Discovery and Delivery* es una revista de acceso abierto que está indexada en Scopus en la categoría de *Library and Information Sciences* está posicionada en un Q2 (25/774) a fecha 19 de junio de 2019. Asimismo, tiene un factor de impacto de 0.34 y un H index de 15.

ISSN: 23986247 en United Kingdom.

Estándares de calidad del artículo: Citas, Vistas, Descargas es un artículo de reciente publicación por lo que aún no se han registrado.

(e) [Artículo 5](#)

Sáiz, M.C., Marticorena, R., Arnaiz-González, Á., & Díez-Pastor, J.F., García-Osorio, C.I. (en prensa). Measuring functional abilities in children (3-6 years old):

combining observational method with computer tool. *Jove-Journal of Visualized Experiments*.

Estándares de calidad de la publicación: es una revista de acceso abierto que está indexada en la Web of Science en la categoría de Multidisciplinary Sciences y Multidisciplinary en SSCI. La revista en la categoría Multidisciplinary está posicionada en un Q1 (9/52) y en un Q2 (31/64) en la categoría de Multidisciplinary Sciences (57/135). Asimismo, tiene un factor de impacto de 1.184 . Además, tiene un percentil 52.344. Datos extraídos a fecha 19 de junio de 2019.

ISSN: 1940-087X.

Estándares de calidad del artículo: Citas, Vistas, Descargas Actualmente, está en prensa.

2.2. Publicaciones secundarias

Se presentan como méritos secundarios de esta tesis doctoral: 2 artículos indexados en CIRC, 7 capítulos de libro, 3 libros y 5 Proceeding 3 de ellos indexados en ISI Conference Proceeding Citation Index y una aplicación informática registrada «e-EarlyCare: Desarrollo de una Aplicación informática para el registro y automatización de la corrección de una Escala de detección de Habilidades Funcionales para niños 0-6 años» realizada con las ayudas para la transferencia del conocimiento Prueba Concepto VI Edición: Impulso a la Valorización y Comercialización de Resultados de la Universidad de Burgos financiada con PLAN TCUE 2018-2020, Cofinanciada por el Fondo Europeo de Desarrollo Regional (FEDER) y la Junta de Castilla y León.

Artículos (2)

(1) [Artículo 6](#)

Sáiz, M.C., Marticorena, R., Arnaiz-González, Á., Escolar, M.C., & Queiruga-Dios, M.Á. (2018). Detección del alumno en riesgo en titulaciones de Ciencias de la Salud: aplicación de técnicas de Learning Analytics. *European Journal of Investigation in Health, Psychology and Education*, 8(3), 129-142. doi: 10.30552/ejihpe.v8i3.273.

Estándares de calidad de la publicación: es una revista de acceso abierto que está indexada en Emerging Sources Citation Index, DICE, ISOC y CIRC, entre otras en la categoría de Social Science, Health Science y Tecnología. ISSN: 2174-8144 eISSN: 2254-9625

Estándares de calidad del artículo: *Citas, Vistas, Descargas* es un artículo de reciente publicación y aún no se han registrado citas.

(2) [Artículo 7](#)

Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., & Arnaiz-González, Á. (2018). Cuestionarios de e-autoevaluación y e-feedback: una aplicación en Moodle. *European Journal of Health Research*, 4(3), 135-148. doi: 10.30552/ejhr.v4i3.116.

Estándares de calidad de la publicación: es una revista de acceso abierto que está indexada en ISOC, CIRC y MEDES, entre otras en la categoría de Health Science y Tecnología. ISSN: 2444-9067 eISSN: 2445-0308

Estándares de calidad del artículo: Citas, Vistas, Descargas es un artículo de reciente publicación y aún no se han registrado citas.

(3) **Artículo 8**

Sáiz, M.C., & Arnaiz-González, Á. (2017). Más allá del Flipped Classroom: Una experiencia de Flipped Learning en la Universidad de Burgos. *Revista Flipped classroom*. Recuperado de <https://www.theflippedclassroom.es/experiencias-flipped-30-mas-alla-del-flipped-classroom-una-experiencia-de-flipped-learning-en-la-universidad-de-burgos-maria-consuelo-saiz-manzanares/>

Capítulos de libro (7)

(1) **Capítulo 1.** Sáiz, M.C., Escolar, M.C., Marticorena, R., García-Osorio, C.I., & Queiruga, M.A. (2017). Formación del profesorado en Metodologías Activas desde Plataformas interactivas. En J.C Núñez., et al. *Temas actuales de investigación en las áreas de la Salud y la Educación* (pp. 39-44). Oviedo: SCINFOPER. Recuperado de https://www.researchgate.net/publication/327069368_Formacion_del_profesorado_en_Metodologias_Activas_desde_Plataformas_interactivas

(2) **Capítulo 2.** Sáiz, M.C., Queiruga, M.A., Marticorena, R., García-Osorio, C.I., & Escolar, M.C. (2017). Análisis de protocolos de pensar en voz alta: un ejemplo de SRL en el aprendizaje de la física. En J.C Núñez., et al. *Temas actuales de investigación en las áreas de la Salud y la Educación* (pp. 731-738). Oviedo: SCINFOPER. Recuperado de https://www.researchgate.net/publication/327121292_Analisis_de_protocolos_de_pensar_en_voz_alta_un_ejemplo_de_SRL_en_el_aprendizaje_de_la_fisica

(3) **Capítulo 3.** Sáiz., M.C., Escolar, M.C., Marticorena, R., García-Osorio, C.I., & Queiruga, M.A. (2017). Aprendizaje basado en proyectos utilizando LMS: una experiencia en Ciencias de la Salud. En J.C Núñez., et al. *Temas actuales de investigación en las áreas de la Salud y la Educación* (pp. 739-746). Oviedo: SCINFOPER. Recuperado de https://www.researchgate.net/publication/327069491_Aprendizaje_basado_en_proyectos_utilizando_LMS_una_experiencia_en_Ciencias_de_la_Salud

(4) **Capítulo 4.** Sáiz., M.C., Escolar, M.C., Marticorena, R., García-Osorio, C.I., & Queiruga, M.A. Conductas de aprendizaje en LMS, SRL y *feedback* efectivo en *B-Learning*. En J.C Núñez., et al. *Temas actuales de investigación en las áreas de la Salud y la Educación*

(pp. 747-752). Oviedo: SCINFOPER. Recuperado de https://www.researchgate.net/publication/327123793_Conductas_de_aprendizaje_en_LMS_SRL_y_feedback_efectivo_en_B-Learning

(5) **Capítulo 5.** Sáiz., M.C., Marticorena, R., Arnaiz, Á., Escolar, M.C., & Queiruga, M.A. (2018). *Flipped Learning* en titulaciones de salud: un acercamiento a la tutorización inteligente. En M. Molero et al. *Intervención en Contextos Clínicos y de la Salud* (Vol. II, pp. 235-242). Almería: ASUNIVEP. Recuperado de https://www.researchgate.net/publication/329844657_Flipped_Learning_en_titulaciones_de_salud_Un_acercamiento_a_la_tutorizacion_inteligente

(6) **Capítulo 6.** Marticorena, R., Sáiz., M.C., Arnaiz, Á., Escolar, M.C., & Queiruga, M.A. (2018). Análisis de los resultados de aprendizaje en Ciencias de la Salud: *Learning Analytics* desde un *plugin* para Moodle. En M. Molero et al. *Intervención en Contextos Clínicos y de la Salud* (Vol. II, pp. 243-252). Almería: ASUNIVEP. Recuperado de https://www.researchgate.net/publication/329844746_Analisis_de_los_resultados_de_aprendizaje_en_Ciencias_de_la_Salud_Learning_Analytics_desde_un_plugin_para_Moodle

(7) **Capítulo 7.** Escolar, M.C., Sáiz., M.C., Arnaiz, Á., Queiruga, M.A., & Marticorena, R. Relación entre los conocimientos previos de los estudiantes de Ciencias de la Salud y las respuestas de aprendizaje en experiencias *Flipped Classroom*. En M. Molero et al. *Intervención en Contextos Clínicos y de la Salud* (Vol. II, pp. 261-270). Almería: ASUNIVEP. Recuperado de https://www.researchgate.net/publication/329844669_Relacion_entre_los_conocimientos_previos_de_los_estudiantes_de_Ciencias_de_la_Salud_y_las_respuestas_de_aprendizaje_en_experiencias_Flipped_Classroom

Libros (3)

Además, se han publicado dos libros por el Servicio de Publicaciones de la Universidad de Burgos referidos al diseño de la plataforma Moodle siguiendo la metodología de aprendizaje metacognitivo y proponiendo una estructura de e-orientación en Educación Superior.

(1) **Libro 1.** Sáiz, M.C., & Montero, E. (2016). *Metodologías activas en docencia universitaria. Diseño de una asignatura de ciencias de la salud en la plataforma virtual.*

[*Active methodologies in university teaching. Design of a health sciences study module on the virtual platform*]. Burgos: Servicio de Publicaciones de la Universidad de Burgos. <https://www.ubu.es/catalogo-de-publicaciones/metodologias-activas-en-docencia-universitaria-diseno-de-una-asignatura-de-ciencias-de-la-salud-en-la-plataforma-virtual>

Y un libro en el que se abordan aspectos de la utilización de técnicas de minería de datos para la investigación en contextos de salud.

(2) **Libro 2.** Sáiz, M.C. (2018). *E-Project Based Learning en Terapia Ocupacional: Una aplicación en la asignatura de «Estimulación Temprana»*. Burgos: Servicio de Publicaciones de la Universidad de Burgos. <https://www.ubu.es/catalogo-de-publicaciones/e-project-based-learning-en-terapia-ocupacional-una-aplicacion-en-la-asignatura-estimulacion-temprana>

(3) **Libro 3.** Sáiz, M.C., Escolar, M.C., & Rodríguez-Medina, J. (en prensa). *Diseño y aplicación de la investigación cualitativa: aplicación de la metodología observacional de métodos mixtos y de técnicas de minería de datos*. Burgos: Servicio de Publicaciones de la Universidad de Burgos.

Proceeding (5)

(1) **Proceeding 1.** Sáiz, M.C., García-Osorio, C.I., & Marticorena, R. (2017, junio). Programa individualizado de e-orientación al estudiante desde la plataforma UbuVirtual: aplicación de técnicas de minería de datos. En V. Abella., V. Ausín, & V. Delgado (Eds.), *Actas XXV Jornadas Universitarias de Tecnología Educativa. JUTE Burgos* (484-490). EDTNTEC: Burgos (España).

(2) **Proceeding 2.** Díez-Pastor, J.F., Arnaiz-González, A., Marticorena, R., & Sáiz, M.C. (2017, junio). ¿Puede el registro de actividad en Moodle predecir el éxito o fracaso de los estudiantes? En V. Abella., V. Ausín, & V. Delgado (Eds.), *Actas XXV Jornadas Universitarias de Tecnología Educativa. JUTE Burgos* (637-444). EDTNTEC: Burgos (España).

(3) **Proceeding 3.** Sáiz, M.C., Zaparín, M.J., Rodríguez-Arribas, S., & Bustillo, A. (2019, march). Design of a smart art classroom in art history: a learning experience with Self-regulated serious games. En L. Gómez Chova., A. López-Martínez., & I. Candel-Torres (Eds.), *13th International Technology, Education and Development Conference*

Proceedings of INTED2019 Conference 11th-13th March 2019 (1998-2006). IATED: Valencia (Spain).

(4) [Proceeding 4](#). Sáiz, M.C., Marticorena, R., Arnaiz-González, Á., Díez-Pastor, J.F., & Rodríguez-Arribas, S. (2019). Blended learning: an experience with infographics and virtual laboratories using Self-regulated learning. *13th International Technology, Education and Development Conference Proceedings of INTED2019 Conference 11th-13th March 2019* (2966-2971). IATED: Valencia (Spain).

(5) [Proceeding 5](#). Sáiz, M.C., Marticorena, R., Arnaiz-González, Á., Díez-Pastor, J.F., & Rodríguez-Arribas, S. (2019). Computer application for the registration and automation of the correction of a functional skills detection scale in Early Care. *13th International Technology, Education and Development Conference Proceedings of INTED2019 Conference 11th-13th March 2019* (5322-5328). IATED: Valencia (Spain).

2.3. Otras acciones de difusión de la investigación

Comunicaciones en Congresos/ reuniones científicas

Ponencias invitadas: 18

	Ponencia invitada
Tipo	Ponencia invitada 1
Autores	Sáiz, M.C., & Montero, E
Título	Metodologías activas en la docencia universitaria: aprendizaje Basado en Proyectos integración en UBUVirtual
Lugar	Instituto de Formación e Innovación Educativa Universidad de Burgos
Horas	17 horas
Fecha	del 17 de octubre al 28 de noviembre de 2016

	Ponencia invitada
Tipo	Ponencia invitada 2
Autores	Sáiz, M.C
Título	Metodologías activas de aprendizaje aplicadas a un aula de formación profesional Inicial
Lugar	Confederación de Asociaciones Empresariales de Burgos (FAE) Dirección Provincial de Educación de Burgos
Horas	10 horas
Fecha	14/11/2017 al 28/11/2017 Burgos

	Ponencia invitada
Tipo	Ponencia invitada 3
Autores	Sáiz, M.C
Título	Metodologías activas de aprendizaje aplicadas a un aula de formación profesional Inicial
Lugar	Confederación de Asociaciones Empresariales de Burgos (FAE) Dirección Provincial de Educación de Burgos
Horas	10 horas
Fecha	16/11/2017 al 30/11/2017. Aranda de Duero (Burgos)

	Ponencia invitada
Tipo	Ponencia invitada 4
Autores	Sáiz, M.C., & Marticorena, R
Título	Técnicas innovadoras en Educación: Aprendizaje Basado en Proyectos y Flipped Learning
Lugar	FabLab Burgos
Horas	2 horas
Fecha	Burgos 27 de octubre de 2017

	Ponencia invitada
Tipo	Ponencia invitada 5
Autores	Sáiz, M.C
Título	Task analysis in the teaching process from Learning Management System (LMS) environment
Lugar	Conferencias invitadas a alumnos de doctorado de la Universidad do Minho en Portugal en el Programa de Doctorado de Ciencias de la Educación en el Instituto de Educación
Horas	3 horas
Fecha	Universidad do Minho (Braga) 24 de Julio de 2018

	Ponencia invitada
Tipo	Ponencia invitada 6
Autores	Sáiz, M.C
Título	Learning analytics Techniques for analysis of students' response
Lugar	Conferencias invitadas a alumnos de doctorado de la Universidad do Minho en Portugal en el Programa de Doctorado de Ciencias de la Educación en el Instituto de Educación
Horas	3 horas
Fecha	Universidad do Minho (Braga) 24 de Julio de 2018

	Ponencia invitada
Tipo	Ponencia invitada 7
Autores	Sáiz, M.C
Título	Feedback to the student's tasks from the LMS environment. An analysis of Learning Style in the (LMS) environment
Lugar	Conferencias invitadas a alumnos de doctorado de la Universidad do Minho en Portugal en el Programa de Doctorado de Ciencias de la Educación en el Instituto de Educación
Horas	3 horas
Fecha	Universidad do Minho (Braga) 25 de Julio de 2018

	Ponencia invitada
Tipo	Ponencia invitada 8
Autores	Sáiz, M.C
Título	An application of Data Mining Techniques in Education (neural networks, multiple regression analysis, automatic linear model, decision trees)
Lugar	Conferencias invitadas a alumnos de doctorado de la Universidad do Minho en Portugal en el Programa de Doctorado de Ciencias de la Educación en el Instituto de Educación
Horas	3 horas
Fecha	Universidad do Minho (Braga) 26 de Julio de 2018

	Ponencia invitada
Tipo	Ponencia invitada 9
Autores	Sáiz, M.C
Título	An application of Data Mining Techniques in Education (neural networks, multiple regression analysis, automatic linear model, decision trees)
Lugar	Orientations to design teaching in the LMS environment
Horas	2 horas
Fecha	Universidad do Minho (Braga) 26 de Julio de 2018

	Ponencia invitada
Tipo	Ponencia invitada 10
Autores	Sáiz, M.C
Título	How does Project-Based Learning in interactive platform improve learning-teaching process?
Lugar	Orientations to design teaching in the LMS environment
Horas	2 horas
Fecha	ERASMUS + KA219 Project 2017—ES01-KA219-038352. Amazing Technologies to enhance learning trough innovation, enjoyment and research. Burgos 22 de octubre de 2018

	Ponencia invitada
Tipo	Ponencia invitada 11
Autores	Sáiz, M.C., Marticorena, R., Arnaiz-González, Á.
Título	Introducción a la programación didáctica basada en metodologías activas en Aprendizaje Basado en Proyectos
Lugar	IFIE Burgos
Horas	2 horas
Fecha	Curso de formación para profesores universitarios. <i>Flipped Learning</i> : diseño y aplicación de Aprendizaje Basado en Proyectos y <i>Flipped Classroom</i> en UBUVirtual. 20 de septiembre de 2018

	Ponencia invitada
Tipo	Ponencia invitada 12
Autores	Sáiz, M.C., Marticorena, R., Arnaiz-González, Á.
Título	Elaboración de rúbricas alineación de las competencias a implementar y los criterios de evaluación
Lugar	IFIE Burgos
Horas	2 horas
Fecha	Curso de formación para profesores universitarios. <i>Flipped Learning</i> : diseño y aplicación de Aprendizaje Basado en Proyectos y <i>Flipped Classroom</i> en UBUVirtual. 27 de septiembre de 2018

	Ponencia invitada
Tipo	Ponencia invitada 13
Autores	Sáiz, M.C., Marticorena, R., Arnaiz-González, Á.
Título	Diseño y estructuración de proyectos
Lugar	IFIE Burgos
Horas	2 horas
Fecha	Curso de formación para profesores universitarios. <i>Flipped Learning</i> : diseño y aplicación de Aprendizaje Basado en Proyectos y <i>Flipped Classroom</i> en UBUVirtual. 4 de octubre de 2018

	Ponencia invitada
Tipo	Ponencia invitada 14
Autores	Sáiz, M.C., Marticorena, R., Arnaiz-González, Á.
Título	Utilización de plataformas virtuales
Lugar	IFIE Burgos
Horas	2 horas
Fecha	Curso de formación para profesores universitarios. <i>Flipped Learning</i> : diseño y aplicación de Aprendizaje Basado en Proyectos y <i>Flipped Classroom</i> en UBUVirtual. 17 de octubre de 2018

Tipo	Ponencia invitada 15
Autores	Sáiz, M.C., Marticorena, R., Arnaiz-González, Á.
Título	Presentación de experiencias <i>Flipped Learning</i>
Lugar	IFIE Burgos
Horas	2 horas
Fecha	Curso de formación para profesores universitarios. <i>Flipped Learning</i> : diseño y aplicación de Aprendizaje Basado en Proyectos y <i>Flipped Classroom</i> en UBUVirtual. 24 de octubre de 2018

Tipo	Ponencia invitada 16
Autores	Sáiz, M.C
Título	Metodologías activas de aprendizaje aplicadas al aula de formación profesional Inicial
Lugar	Confederación de Asociaciones Empresariales de Burgos (FAE). Dirección Provincial de Educación de Burgos
Horas	10 horas
Fecha	06/11/2018 al 20/11/2018. Burgos

Tipo	Ponencia invitada 17
Autores	Sáiz, M.C
Título	«Analysis of functional skills in young children with disabilities through application eEarlyCare: a detection and implementation of person-to-person therapy programs»
Lugar	Conferencia invitada en la Universidad do Minho Braga (Portugal)
Horas	4 horas
Fecha	3 de mayo de 2019

Tipo	Ponencia invitada 18
Autores	Sáiz, M.C
Título	«Analysis of functional skills in young children with disabilities through application «Validation of a Scale of Student Satisfaction»
Lugar	Conferencia invitada en la Universidad do Minho Braga (Portugal)
Horas	4 horas
Fecha	4 de mayo de 2019

Workshops: 2

	Workshop 1
Tipo	Workshop
Autores	Sáiz, M.C., & Alegre, J.M
Título	Evaluación y Feedback en Educación Superior
Lugar	Escuela Politécnica Superior Universidad de Burgos. Instituto de Formación e Innovación Educativa. Universidad de Burgos
Horas	4 horas
Fecha	4 de noviembre de 2016

	Workshop 2
Tipo	Workshop
Autores	Sáiz, M.C., & Arnaiz-González, Á
Título	Herramientas y Metodologías de Feedback, Evaluación y Análisis a través de Técnicas de Minería de Datos en entornos de aprendizaje virtual.
Lugar	Escuela Politécnica Superior Universidad de Burgos. Instituto de Formación e Innovación Educativa. Universidad de Burgos
Horas	4 horas
Fecha	4 de noviembre de 2016

Comunicaciones: 27

	Comunicación oral
Tipo	Comunicación oral. 1
Autores	Sáiz, M.C., García-Osorio, C.I., & Marticorena, R.
Título	Programa individualizado de e-orientación al estudiante desde la plataforma Ubuvirtual: aplicación de Técnicas de Minería de Datos
Organismo	JUTE Burgos
Horas	2 horas
Fecha	Burgos 21-13 junio 2017

	Comunicación oral
Tipo	Comunicación oral. 2
Autores	Díez-Pastor, J.F., Arnaiz-González, A., Marticorena, R., & Sáiz, M.C
Título	¿Puede el registro de actividad en Moodle predecir el éxito o fracaso de los estudiantes?
Organismo	JUTE Burgos
Horas	2 horas
Fecha	Burgos 21-13 junio 2017

	Comunicación oral
Tipo	Comunicación oral. 3
Autores	Sáiz, M.C
Título	Aprendizaje Basado en Proyectos en Educación Superior: una experiencia en Ciencias de la Salud
Organismo	III Congreso Nacional de Psicología
Horas	2 horas
Fecha	Oviedo del 3 al 7 de julio de 2017

	Comunicación oral
Tipo	Comunicación oral. 4
Autores	Escolar, M.C., Sáiz, M.C., Marticorena, R., Queiruga, M.Á., & García-Osorio, C.I
Título	Formación del profesorado en metodologías activas en plataformas interactivas.
Organismo	I Congreso Internacional de Psicología, Salud y Educación
Horas	media hora
Fecha	Oviedo del 3 al 7 de julio de 2017

	Comunicación oral
Tipo	Comunicación oral. 5
Autores	Queiruga, M.Á., Sáiz, M.C., Marticorena, R., Escolar, M.C., & García-Osorio, C.I
Título	Análisis de protocolos de pensar en voz alta: un ejemplo en SRL en el aprendizaje de la Física.
Organismo	I Congreso Internacional de Psicología, Salud y Educación
Horas	media hora
Fecha	Oviedo del 3 al 7 de julio de 2017

	Comunicación oral
Tipo	Comunicación oral. 6
Autores	Sáiz, M.C., Escolar, M.C., Marticorena, R., Queiruga, M.Á., & García-Osorio, C.I
Título	Trabajo en aprendizaje basado en proyecto desde una plataforma interactiva: una experiencia en ciencias de la salud.
Organismo	I Congreso Internacional de Psicología, Salud y Educación
Horas	media hora
Fecha	Oviedo del 3 al 7 de julio de 2017

	Comunicación oral
Tipo	Comunicación oral. 7
Autores	Marticorena, R., Sáiz, M.C., Escolar, M.C., Queiruga, M.Á., & García-Osorio, C.I
Título	Conductas de aprendizaje en LMS: SRL y feedback efectivo en <i>Blended-Learning</i>
Organismo	I Congreso Internacional de Psicología, Salud y Educación
Horas	media hora
Fecha	Oviedo del 3 al 7 de julio de 2017

	Comunicación oral
Tipo	Comunicación oral. 8
Autores	Sáiz, M.C., & Arnaiz-González, Á
	Como diseñar un sistema de tutorización interactiva inteligente en las plataforma virtuales: aplicación de estrategias de autorregulación y feedback orientado a procesos. Estudios de caso
Organismo	Instituto de Formación e Innovación Educativa. Universidad de Burgos
Horas	4
Fecha	16 horas del 11 y 12 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 9
Autores	Sáiz, M.C., & Arnaiz-González, Á
	Técnicas de Minería de Datos aplicadas al contexto educativo
Organismo	Instituto de Formación e Innovación Educativa. Universidad de Burgos
Horas	4
Fecha	16 horas del 11 y 12 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 10
Autores	Sáiz, M.C., Arnaiz-González, Á., Marticorena, R., Díez-Pastor, J.F., García-Osorio, C.I., Escolar, M.C
Título	Análisis del trabajo colaborativo en contextos de B-Learning en Educación Superior: aplicación de Educational Data Mining
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Logroño 21-23 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 11
Autores	Sáiz, M.C., Arnaiz-González, Á., Marticorena, R., Díez-Pastor, J.F., García-Osorio, C.I., Escolar, M.C
Título	Estudio del rendimiento académico y de las conductas de aprendizaje en LMS: aplicación de un análisis de clústeres
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Logroño 21-23 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 12
Autores	Sáiz, M.C., Arnaiz-González, Á., Marticorena, R., Díez-Pastor, J.F., García-Osorio, C.I., Escolar, M.C
Título	¿Son los conocimientos previos y las estrategias metacognitivas predictores de aprendizajes eficaces?
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Logroño 21-23 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 13
Autores	Sáiz, M.C., Arnaiz-González, Á., Marticorena, R., Díez-Pastor, J.F., García-Osorio, C.I., Escolar, M.C
Título	Un método automático para la medición del uso de estrategias metacognitivas en análisis de protocolos
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Logroño 21-23 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 14
Autores	Sáiz, M.C., Arnaiz-González, Á., Marticorena, R., Díez-Pastor, J.F., García-Osorio, C.I., Escolar, M.C
Título	Predicción de estudiantes en riesgo: aplicación de técnicas de aprendizaje automático
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Logroño 21-23 de junio de 2018

	Comunicación oral
Tipo	Comunicación oral. 15
Autores	Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., Arnaiz-González, Á.
Título	Cuestionarios de auto-evaluación y feedback: un análisis diferencial de incidencia en Ciencias de la Salud
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Madrid del 21-23 de noviembre de 2018

	Comunicación oral
Tipo	Comunicación oral. 16
Autores	Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., Arnaiz-González, Á.
Título	Visualización de los resultados de aprendizaje: una herramienta de vinculación con Moodle
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Madrid del 21-23 de noviembre de 2018

	Comunicación oral
Tipo	Comunicación oral. 17
Autores	Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., Arnaiz-González, Á.
Título	Aplicación de Técnicas de Learning Analytics para la detección del alumno en riesgo en estudios de Ciencias de la Salud
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Madrid del 21-23 de noviembre de 2018

	Comunicación oral
Tipo	Comunicación oral. 18
Autores	Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., Arnaiz-González, Á.
Título	Relación entre los conocimientos previos de los estudiantes y las respuestas de aprendizaje en experiencias <i>Flipped Classroom</i> en Ciencias de la Salud
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Madrid del 21-23 de noviembre de 2018

	Comunicación oral
Tipo	Comunicación oral. 19
Autores	Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., Arnaiz-González, Á.
Título	Aplicación de Técnicas de <i>Flipped Learning</i> en titulaciones de salud: un estudio comparativo
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	Media hora
Fecha	Madrid del 21-23 de noviembre de 2018

Tipo	Comunicación oral. 20
Autores	Sáiz, M.C., Zaparaín, M.J., Rodríguez-Arribas, S., Bustillo, A
Título	Design of a smart art classroom in art history: a learning experience with Self-regulated serious games
Organismo	INTED 2019
Horas	Media hora
Fecha	11-13 march Valencia (Spain)

Tipo	Comunicación oral. 21
Autores	Sáiz, M.C, Marticorena, R., Rodríguez-Arribas, S
Título	<i>Blended Learning</i> : an experience with infographics and virtual laboratories using Self-regulated learning
Organismo	INTED 2019
Horas	Media hora
Fecha	11-13 march Valencia (Spain)

Tipo	Comunicación oral. 22
Autores	Sáiz, M.C, Rodríguez-Arribas, S., Marticorena, R.
Título	Computer application for the registration and automation of the correction of a functional skills detection scale in Early Care
Organismo	INTED 2019
Horas	Media hora
Fecha	11-13 march Valencia (Spain)

Tipo	Comunicación oral. 23
Autores	Sáiz, M.C, Rodríguez-Arribas, S., Pardo, C.
Título	Aprendizaje Autorregulado en <i>Blockly games</i> : un análisis de la efectividad de implementación en distintos entornos
Organismo	INTED 2019
Horas	Media hora
Fecha	11-13 march Valencia (Spain)

Tipo	Comunicación oral. 24
Autores	Sáiz, M.C, Marticorena, R., Arnaiz-Gonzales, Á
Título	Cómo medir habilidades funcionales en niños pequeños: uso de <i>eEarlyCare computer application</i>
Organismo	XV Congreso Internacional Galego-Portugués de Psicopedagogía
Horas	1 hora
Fecha	4-6 de septiembre A Coruña (Spain)

Tipo	Comunicación oral. 25
Autores	Sáiz, M.C, Marticorena, R., Antón-Caballero, M., Velasco-Saiz, R
Título	Programas de intervención temprana en niños pequeños: uso de <i>eEarlyCare Program</i>
Organismo	XV Congreso Internacional Galego-Portugués de Psicopedagogía
Horas	1 hora
Fecha	4-6 de septiembre A Coruña (Spain)

Tipo	Comunicación oral. 26
Autores	Sáiz, M.C, Marticorena, R., Zaparaín, M.J., Arnaiz-Gonzales, Á., Velasco-Saiz, R
Título	Análisis de tareas a través de eye tracking en Historia del Arte
Organismo	XV Congreso Internacional Galego-Portugués de Psicopedagogía
Horas	1 hora
Fecha	4-6 de septiembre A Coruña (Spain)

Tipo	Comunicación oral. 27
Autores	Sáiz, M.C, Zaparaín, M.J., Marticorena, R.,
Título	Aprendizaje Autorregulado a través de la gamificación: una propuesta de innovación docente en Historia del Arte
Organismo	XV Congreso Internacional Galego-Portugués de Psicopedagogía
Horas	1 hora
Fecha	4-6 de septiembre A Coruña (Spain)

Coordinación de Simposios: 4

	Coordinación de Simposio
Tipo	Coordinación de Simposio. 1
Autores	Sáiz, M.C
Título	Coordinadora del Simposio invitado «Enseñanza-aprendizaje en entornos de docencia virtual: ¿cómo autorregular y retroalimentar desde las plataformas virtuales?»
Organismo	I Congreso Internacional de Psicología, Salud y Educación
Horas	1 horas
Fecha	Oviedo 8-11 de noviembre de 2017

	Coordinación de Simposio
Tipo	Coordinación de Simposio. 2
Autores	Sáiz, M.C., Arnaiz-González, Á., Marticorena, R., Díez-Pastor, J.F., García-Osorio, C.I., Escolar, M.C
Título	Coordinadora del Simposio invitado: Aprendizaje personalizado en Educación Superior: aplicación de técnicas de inteligencia artificial y de minería de datos https://www.ubu.es/data-analysis-techniques-applied-health-environments-sciences-datahes-0/por-un-aprendizaje-personalizado
Organismo	IX Congreso Internacional de Psicología y Educación
Horas	1 hora
Fecha	Logroño 21-23 de junio de 2018

	Coordinación de Simposio
Tipo	Coordinación de Simposio. 3
Autores	Sáiz, M.C., Queiruga-Dios, M.Á., Marticorena, R., Escolar, M.C., Arnaiz-González, Á.
Título	Coordinadora del Simposio invitado: Técnicas de enseñanza, aprendizaje y tecnología en el Siglo XXI: nuevos horizontes en Ciencias de la Salud en Educación Superior https://www.ubu.es/noticias/la-ubu-la-cabeza-en-la-investigacion-sobre-aprendizaje-del-siglo-xxi
Organismo	V Congreso Internacional en Contextos Psicológicos, Educativos y de la Salud
Horas	3 horas
Fecha	Madrid del 21-23 de noviembre de 2018

	Coordinación de Simposio
Tipo	Coordinación de Simposio. 4
Autores	Sáiz, M.C.
Título	Coordinadora del Simposio invitado: Metacognición, autorregulación y evaluación a través de <i>Cloud Technology</i> , <i>Artificial Intelligence</i> y <i>Educational Data Mining</i> https://congreso-xvgp.asocip.com/index.php/es/
Organismo	XV Congreso Internacional Galego-Portugués de Psicopedagogía
Horas	3 horas
Fecha	A Coruña del 4 al 6 de septiembre de 2019

Poster: 3

	Póster
Tipo	Póster 1
Autores	Velasco, R., Pérez, Y., Sáiz, M.C., Marticorena, R
Título	Desarrollo de una aplicación para el registro e interpretación en una escala de análisis funcional para niños 0-6 años
Organismo	V Congreso Internacional en Contextos Psicológicos, Educativos y de la Salud
Horas	2 horas
Fecha	Madrid del 21-23 de noviembre de 2018

	Póster
Tipo	Póster 2
Autores	Velasco-Sáiz, R., Sáiz, M.C., Campo-Roca, V.
Título	¿Existe relación entre las habilidades metalingüísticas, metarrepresentacionales y metacognitivas en niños de tres años?
Organismo	V Congreso Internacional en Contextos Psicológicos, Educativos y de la Salud
Horas	2 horas
Fecha	Madrid del 21-23 de noviembre de 2018

	Póster
Tipo	Póster 3
Autores	Sáiz, M.C., Marticorena, R., Arnaiz-González, Á., Díez-Pastor, J.F., Velasco-Sáiz, R.,
Título	eEarlyCare: detección e intervención en habilidades funcionales en niños pequeños a través de una aplicación informática
Organismo	V Congreso Internacional en Contextos Psicológicos, Educativos y de la Salud
Horas	2 horas
Fecha	Madrid del 21-23 de noviembre de 2018

Aplicación informática

Contiene un sistema de registro y de análisis de datos para ser trasferidos y analizados en paquetes estadísticos como SPSS, WEKA, o librerías de Python que permitirán la aplicación de Técnicas de Minería de Datos. <https://www.ubu.es/noticias/la-ubu-mide-las-habilidades-de-los-ninos-de-0-6-anos>

Sáiz, M.C., Marticorena, R., Arnaiz-González., & Díez-Pastor, J.F. eEarly Care BU-07-19. Número de asiento registral 00/2019/1663

2.4. Relación de las publicaciones y las técnicas de estadísticas y de minería de datos implementadas

2.4.1. Publicaciones principales

Publicación	Técnicas estadísticas	Técnicas de Minería de Datos
Artículo 1	Análisis de asimetría y de curtosis Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto eta cuadrado Test de Bonferroni	Análisis de clúster con el algoritmo <i>k</i> -means Análisis discriminante

Publicación	Técnicas estadísticas	Técnicas de Minería de Datos
Artículo 2	Matriz de correlación (coeficiente de correlación de Pearson) Prueba U de Mann-Whitney y Test de Wilcoxon Análisis de asimetría y de curtosis Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto eta cuadrado Test de Bonferroni CrossTable	Análisis de clúster con el algoritmo <i>k</i> -means Wilk's Lambda Función canónica discriminante
Artículo 3	Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto eta cuadrado Test de Bonferroni ANCOVAS	
Artículo 4	Estadísticos descriptivos Prueba U de Mann-Whitney Percentiles Gráficos de base radial	Análisis de clúster con el algoritmo <i>k</i> -means
Artículo 5	Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto eta cuadrado Test de Bonferroni	Árbol de decisión Redes Neuronales de Base Radial Regresión Lineal Perceptrón Multicapa Support Vector Machine <i>Knn</i> de clasificación <i>ibk</i> (vecino más cercano)

2.4.2. Publicaciones secundarias

Publicación	Técnicas estadísticas	Técnicas de Minería de Datos
Artículo 6	Estadísticos descriptivos Alpha de Cronbach Matriz de correlaciones (coeficiente de correlación de Pearson)	Regresión Lineal Automática Análisis de clúster con el algoritmo <i>k</i> -means Gráficos de regresión general y parcial

Publicación	Técnicas estadísticas	Técnicas de Minería de Datos
Artículo 7	ANOVA de un factor de efectos fijos Gráficos de base radial Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto eta cuadrado	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Artículo 8	ANCOVAS Estadísticos descriptivos	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Capítulo 1	Estadísticos descriptivos Coeficiente de Variación Alpha de Cronbach	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Capítulo 2	Estadísticos descriptivos Prueba de Wilcoxon Valor del efecto <i>d</i> de Cohen Matriz de correlaciones (coeficiente de correlación de Pearson)	
Capítulo 3	Estadísticos descriptivos Alpha de Cronbach	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Capítulo 4	Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto Eta cuadrado	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Capítulo 5	Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto Eta cuadrado	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Capítulo 6	Estadísticos descriptivos Distintos tipos de gráficos implementados a través de una aplicación plugin para Moodle	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual

Publicación	Técnicas estadísticas	Técnicas de Minería de Datos
Capítulo 7	Estadísticos descriptivos	Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Libro 1		Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Libro 2		Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Libro 3		
Proceeding 1	Estadísticos descriptivos ANOVA de un factor de efectos fijos Valor del efecto eta cuadrado	Técnicas de regresión
Proceeding 2		Árboles de decisión algoritmo CART Regresión logística Predicciones usando validación cruzada con <i>Scikit-Learn</i>
Proceeding 3		Técnicas de <i>Learning Analytics</i> en un <i>SmartArt</i> con serious games
Proceeding 4		Técnicas de <i>Learning Analytics</i> en la plataforma basada en Moodle UBUVirtual
Proceeding 5		Técnicas de <i>Learning Analytics</i> en una aplicación informática elaborada <i>ad hoc</i>

RELACIÓN DE FIGURAS Y DE TABLAS

Relación de Figuras y de Tablas. Punto 1. Introducción

Figuras

Figura 1.....	8
Figura 2.....	9
Figura 3.....	12
Figura 4.....	13
Figura 5.....	15
Figura 6.....	25
Figura 7.....	26
Figura 8.....	51
Figura 9.....	52
Figura 10.....	53
Figura 11.....	53
Figura 12.....	54
Figura 13.....	54
Figura 14.....	55
Figura 15.....	55
Figura 16.....	57
Figura 17.....	58
Figura 18.....	59
Figura 19.....	60
Figura 20.....	60
Figura 21.....	62
Figura 22.....	64
Figura 23.....	65
Figura 24.....	65
Figura 25.....	66
Figura 26.....	67
Figura 27.....	69
Figura 28.....	73
Figura 29.....	75
Figura 30.....	76
Figura 31.....	77
Figura 32.....	78
Figura 33.....	79
Figura 34.....	80
Figura 35.....	81
Figura 36.....	81
Figura 37.....	82
Figura 38.....	83
Figura 39.....	83

Figura 40	86
Figura 41	86
Figura 42	86
Figura 43	86
Figura 44	86
Figura 45	87
Figura 46	89
Figura 47	90
Figura 48	90
Figura 49	91
Figura 50	92
Figura 51	93
Figura 52	93
Figura 53	94

Tablas

Tabla 1	16
Tabla 2	28
Tabla 3	41
Tabla 4	50

GLOSARIO DE TÉRMINOS

A

Aprendizaje automático

se puede considerar un subcampo de las ciencias de la computación y una rama de la inteligencia artificial. El objetivo de esta disciplina es desarrollar computadoras que aprendan a partir de la resolución de un problema., 1, 3

Aprendizaje Basado en Proyectos

es un método de aprendizaje que se fundamenta en el trabajo colaborativo desde la resolución de una tarea, problema o proyecto. Los estudiantes para resolverla tienen que aplicar los conocimientos declarativos y los procedimentales. Se utiliza con mucha frecuencia en la rama de ciencias de la salud y en la de ingeniería, 3, 25, 30, 125, 144

Aprendizaje Basado en Proyectos (ABP)

es un método de aprendizaje que se fundamenta en el trabajo colaborativo desde la resolución de una tarea, problema o proyecto. Los estudiantes para resolverla tienen que aplicar los conocimientos declarativos y los procedimentales. Se utiliza con mucha frecuencia en la rama de ciencias de la salud y en la de ingeniería, 25

Aprendizaje no supervisado

Hace referencia al estudio de los patrones de comportamiento, ya que no se tiene información sobre las categorías, 3, 67, 68, 71, 84

Aprendizaje personalizado

consiste en adaptar los conocimientos conceptuales y procedimentales al nivel de conocimientos previos del aprendiz, así como a su forma de procesar la información, 3, 39, 56

Aprendizaje supervisado

es un tipo de aprendizaje automático, que hace referencia a la utilización de algoritmos de clasificación, 3, 33, 68, 70, 71, 75

Artificial Neural Net Works

se entiende como una conexión de sistemas de computación que se inspiran en el comportamiento de las neuronas del cerebro humano. las redes neuronales se basan en algoritmos que trabajan juntos y que procesan entradas de datos complejas, 35

C

CBLE

Computer Based learning environment, son espacios de aprendizaje en plataformas interactivas, 20, 21, 22

Conocimiento declarativo

hace referencia al conocimiento conceptual, 9, 14, 22

Conocimiento metacognitivo

grado de reflexión que tiene un sujeto sobre sus propias habilidades metacognitivas, 7, 8, 10, 12, 14

Conocimiento procedimental

hace referencia al conocimiento sobre el procedimiento, 10, 14

CSCL

Computer Supported Collaborative Learning, son espacios de trabajo colaborativo en plataformas interactivas de aprendizaje, 20, 21, 27, 194, 195

E

Evaluación

estrategias metacognitivas de evaluación del proceso y del producto durante la resolución de una tarea, 9, 16, 20, 50, 147, 153, 154

Experiencias metacognitivas

hace referencia a la experiencia del sujeto en la resolución de tareas semejantes, 7, 10

I

Insight

es un término que se utiliza en psicología y que implica una toma de conciencia sobre un proceso se relaciona con el aprendizaje profundo, 17, 19

Instance-based learning

En el contexto de aprendizaje automático (machine learning), el aprendizaje basado en instancias es un conjunto de algoritmos que compara nuevas instancias de problemas con las instancias vistas en el entrenamiento y que tienen almacenadas en la memoria. Lo que implica que la complejidad de las hipótesis puede crecer con los datos., 35

L

Learning Analytics

Supone la recopilación, análisis, y realización de informes sobre los procesos de aprendizaje de los estudiantes, está relacionado con la minería de datos en educación, 3, 30, 107, 183, 191

Learning Management System

se pueden entender como gestores de aprendizaje, son software instalados en un servidor web que se utiliza para administrar, distribuir y analizar la interacción entre el profesor y los estudiantes y la interacción entre los estudiantes con los distintos recursos para el aprendizaje dispuestos en el sistema, 3, 21, 38

Logs

terminología informática que hace referencia al registro a la grabación secuencial en un archivo o en una

base de datos, 3, 30, 36, 37, 38, 50, 51, 56, 125, 193, 194

M

Método sistemático de enseñanza

hace referencia a una programación didáctica con un hilo conductor desde el objeto de aprendizaje a la consecución del producto de aprendizaje incluido el procedimiento, 26

Monitorización

estrategias metacognitivas de observación antes, durante y después de la resolución de tareas, 8

O

Orientación

estrategias metacognitivas de orientación hacia la resolución de una tarea, 8, 16, 185, 200

P

Planificación

estrategias metacognitivas de planificación de las estrategias de resolución de una tarea, 8, 16, 20, 22, 153, 154, 161, 162

Profundidad en el aprendizaje

deep learning, hace referencia a un aprendizaje reflexionado, autónomo y sólido respecto del concepto aprendido, 9

Python, 1, 3, 51, 119

Es un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, usa tipado dinámico y es multiplataforma, 1

R

Recursividad del pensamiento

la recursividad del pensamiento se puede entender como la capacidad de

pensar sobre el propio pensamiento,
19

Recursos multimedia

son aquellos que utilizan múltiples medios digitales para presentar la información objeto de aprendizaje (texto, imágenes, vídeos, sonido, etc.), 28

Regresión logística

es un tipo de análisis de regresión que se utiliza para predecir el resultado de una variable categórica en función de variables predictoras. Es la probabilidad de predicción de un evento en función de otros., 78, 79

Regulación escalonada

se entiende una jerarquía en orden de importancia desde la técnica de aproximaciones sucesivas a la meta, 9

S

Self-efficacy

termino acuñado por Albert Bandura, que significa el conocimiento de la propia eficacia, 7, 18, 19, 202

T

Técnicas de Minería de Datos

se puede entender como un campo de la estadística o de las ciencias de la computación que pretende hallar patrones en grandes conjuntos de datos. Utiliza métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos, 1, 30, 200

U

UBUVirtual

Es un Learning Management System basado en Moodle que utiliza la Universidad de Burgos como plataforma interactiva de aprendizaje, 37

W

WEKA

Waikato Environment for Knowledge Analysis, es decir «entorno para análisis del conocimiento de la Universidad de Waikato». Es una plataforma libre de software para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato, 1, 3, 51, 70

RELACIÓN DE ACRÓNIMOS

<i>Analíticas de aprendizaje (LA)</i>	30	<i>Knowledge Discovery in Databases (KDD)</i>	62
<i>Aprendizaje Basado en Proyectos (ABP)</i>	30, 125, 144	<i>Learning Management System (LMS)</i> ...	3, 38, 110
<i>Auto-regulación</i> ...	7, 8, 11, 15, 18, 19, 21, 58, 144, 148	<i>Maximización (M)</i>	86
<i>Blended Learning</i> ...	1, 35, 74, 89, 92, 116	<i>Medias (M)</i>	32
<i>Blended Learning (B-Learning)</i>	35	<i>Metacognición</i> 1, 3, 7, 8, 9, 10, 16, 17, 19, 23	
<i>Data Mining (DM)</i>	60	<i>Meta-razonamiento (MT)</i>	24
<i>Desviaciones típicas (DT)</i>	32	<i>Minería de datos educativos (EMD)</i>	30
<i>DSVL (Domain-Specific Visual Language)</i>	23	<i>MOF (Modelo de Organización de funciones)</i>	23
<i>EMD</i>	30, 35, 36, 50, 70, 183	<i>Network analysis (SNA)</i>	70
<i>Escala de estrategias de aprendizaje (ACRAR)</i>	143, 144	<i>Pattern Mining (PM)</i>	60
<i>Expectación-maximización (EM)</i>	86	<i>Process mining (ProM)</i>	69
<i>Face to Face (F2F)</i>	19, 35, 145	<i>ProM</i>	69, 70, 71
<i>Feedback</i> ..	3, 7, 16, 17, 18, 19, 20, 21, 26, 28, 29, 32, 35, 37, 39, 56, 58, 61, 80, 106, 137, 143, 144, 145, 146, 147, 149, 150, 152, 154, 155, 156, 163, 169, 179, 180, 181, 183, 192, 193, 197, 200	<i>Recursos hipermedia</i> ...	21, 23, 39, 57, 58, 61
<i>Flipped Classroom (FL)</i>	57	<i>Recursos hipermedia</i>	57
<i>Flipped Learning (FL)</i> ..	57, 126, 158, 205	<i>Redes Neuronales (RNA)</i>	91
<i>Función de Base Radial (FBR)</i>	91	<i>Redes neuronales de base radial (RNBR)</i>	91
<i>Grado en Enfermería (GE)</i>	129	<i>Replacement Blend (RB)</i>	128, 181
<i>Grado en Terapia Ocupacional (GTO)</i>	129	<i>Replacement blended (RB)</i>	35
<i>Graph Mining (GM)</i>	70	<i>Self Regulated Learning (SRL)</i>	58
<i>Grupo Control (GC)</i>	174	<i>Sequence Pattern Mining (SPM)</i>	70
<i>Grupo Experimental (GE)</i>	174	<i>Supplemental Blend (SB)</i>	181
<i>Intention Mining (IM)</i>	70	<i>Supplemental blended (SB)</i>	35
<i>k-nearest neighbor classification (KNN)</i>	33	<i>Support vector machines (SVM)</i>	33
		<i>Tareas de razonamiento (TR)</i>	24
		<i>Valor de Inflación de la Varianza (VIF)</i>	77
		<i>WEKA</i>	1, 29, 51, 119, 202

