

Minería de datos educativos en plataformas virtuales de aprendizaje musical

Educational Data Mining with Learning Management Systems

Manuel Jesús Espigares Pinazo
IES Profesor Juan Bautista (El Viso del Alcor-Sevilla)
manuelespigares@yahoo.es

Rafael García Pérez
Dto. Métodos de Investigación
y Diagnóstico en Educación
Universidad de Sevilla
rafaelgarcia@us.es

Recibido: 15-12-2010 Aceptado: 7-4-2011

Resumen

El siguiente trabajo presenta la evaluación del proceso de uso de una plataforma virtual de aprendizaje diseñada ad hoc para la Educación Musical en la ESO. Nuestro objetivo se centra en evaluar el uso de dicha plataforma mediante las técnicas de la Minería de datos educativos o Educational Data Mining (EDM). Estas técnicas nos permiten conocer el comportamiento de los datos durante los procesos de enseñanza-aprendizaje en entornos de educación online, lo cual brinda la posibilidad de monitorizar actividades para optimizar su funcionamiento, corregir deficiencias en el aprendizaje, detectar el comportamiento anómalo de determinados grupos de actividad (outliers) y mejorar el diseño de las actividades de enseñanza-aprendizaje online. Las técnicas de análisis que empleamos en la investigación consisten en la limpieza de la información y selección de las variables significativas de estudio, empleo de análisis de conglomerados o tipo cluster, elaboración de gráficas de mallas y reglas predictivas a partir del comportamiento de los datos registrados. Los resultados que obtenemos nos muestran que las actividades que más registros han acumulado han sido de tipo de exploratorio como “ver tema de debate” y “ver foro de discusión”. El alumnado ha empleado la mayor del tiempo en explorar la herramienta. La actividad más intensa se ha generado en las herramientas de comunicación como los foros y los chats donde se han reflejado la mayor parte de los aprendizajes musicales del alumnado.

Palabras clave: Plataformas de Aprendizaje Virtual, Minería de Datos Educativos, Moodle.

Abstract

This work presents the evaluation of the process of using a virtual learning platform designed ad hoc for Music Education in Secondary. Our goal is to analyze the use of this platform through the techniques of Educational Data Mining (EDM) so that we study the application of data mining to data collected with Learning Management System (LMS). These techniques allow us to understand the behavior of the data during the process of teaching and learning in online learning environments, which provides the ability to monitor activities to optimize their performance, correct learning deficiencies, detect anomalous behavior of certain groups of activity (outliers) and improve the design of online teaching-learning activities. The analytical techniques we use in research are oriented to clean the information and selection of significant variables in the study and to use cluster analysis and cluster type and predictive rules from the behavior of the data collected. The results we obtain show that the activities that have accumulated the most records have been exploratory types: “to see the subject of discussion” and “see the discussion forum”; the student body has spent most of time exploring those tools. The most intense activity has been generated by communication tools such as forums and chats where we have recorded the majority of students’ musical learning.

Key words: Learning Management Systems, Educational Data Mining, Moodle.

1. Introducción

Este artículo aborda una nueva metodología de evaluación aplicable en los procesos educativos más característicos de la Sociedad del Conocimiento. Esto es, el aprendizaje regulado mediante plataformas de teleformación (Learning Management System, LMS), que se sitúa en esta sociedad en un primer plano. El empleo de plataformas telemáticas para la formación musical, forma parte de los nuevos rumbos que emprende la ciencia, ampliando y diversificando los enfoques de la Educación Musical (basado en De Aguilera y Adell, 2010; Area, 2004; De Pablos, 1996). Nos referimos a las técnicas estadísticas de la minería de datos (Bin-shyan, Bente y Tsong-wuu, 2008; Chien-ming, Cheng-hao, Romero, Ventura y García, 2008; Espigares y García, 2010; Cox, 2009; Sierra, 2006), que aplicadas en este campo se denominan Minería de Datos Educativos (Educational Data Mining –EDM-). Éstas, implican la generación de modelos estadísticos de carácter correlacional y predictivo (mediante técnicas exploratorias y confirmatorias) con los datos almacenados (registros de las acciones) en dichas plataformas. En esta línea, nuestro trabajo plantea una manera organizada de evaluar distintos objetos de aprendizaje online y dinámicas del proceso de aprendizaje registrado en los LMS.

El EDM, lo aplicamos al estudio de un *modelo educativo-musical* que emplea una plataforma de teleaprendizaje denominada Moodle¹ (ver. 6.0) en centros TIC de secundaria (basado en Cole, 1999; Daniels, 2003; Espigares y García Pérez, 2006). La investigación que llevamos a cabo comprende técnicas de muestreo estadístico que permite estudiar los cerca de 50.000 registros del proceso de uso del sistema de gestión del conocimiento que estudiamos. Esto posibilita el conocimiento de la *frecuencia de uso de los distintos módulos de aprendizaje del sistema* y detectar las *deficiencias del funcionamiento del mismo*. El estudio pormenorizado de los registros de uso de la plataforma facilita tareas como la *monitorización de actividades digitales* de enseñanza-aprendizaje en contextos de teleformación. De este modo, se potencia un mayor control de los procesos formativos y una mejor evaluación de los mismos.

A *nivel teórico*, nuestro modelo educativo, denominado *Modelo Bordón*, recrea una metáfora a partir de la cuerda grave (bordón) de un instrumento musical cordófono, que presenta dos niveles: 1) *El nivel externo* y que recubre la cuerda, representa en nuestra metáfora al *Enfoque Sociocultural*. 2) *El nivel interno*, se compone (como nuestro modelo) de una serie de finos hilos que se entrecruzan y entrelazan dando coherencia y utilidad al modelo educativo-

¹ www.moodle.org.

musical basado en el Enfoque Sociocultural. Ese nivel interno lo constituyen constructos e ideas basadas en el Enfoque Sociocultural. En el siguiente gráfico mostramos un bordón musical, donde observamos los dos niveles teóricos que inspiran nuestro modelo educativo-musical.



Imagen 1. Bordón Musical.

Los hilos del nivel interno suponen, en nuestra metáfora educativa-musical, las 10 claves didácticas del Modelo Bordón sociocultural: 1) Aprendizaje constructivo orientado didácticamente por la indagación y el descubrimiento. 2) Aprendizaje activo fundamentado en la teoría de la actividad. 3) Mediación sociocultural a través de artefactos ideales/materiales (TIC del aprendizaje digital y del proceso creativo musical). 4) División y organización grupal del trabajo, autonomía del aprendizaje y autogestión. 5) Resolución de problemas y conflictos híbridos planteados como oportunidad de aprendizaje. 6) Interacciones sociales orientadas al entendimiento y al trabajo colaborativo en el aprendizaje. 7) Enseñanza y evaluación recíprocas entre el propio alumnado, tomando roles de apoyo, tutorización y seguimiento de los aprendizajes colaborativos de los/as compañeros/as para su evaluación formal (con la intervención del profesorado y mediación de la plataforma). 8) Aproximación e hibridación entre el entorno vital cotidiano y el entorno escolar formal. 9) Comunicación, discurso y argumentación con distintas modalidades del lenguaje implicadas en las actividades: cara a cara, textual formal, chats, foros... 10) Integración sociocultural e interacción didáctica de todos los grupos de alumnado (de los 4 niveles de secundaria) en la plataforma de teleformación como comunidad de aprendizaje global e inclusiva para todas las personas que participan en la experiencia.

En nuestro modelo pedagógico Bordón, mediante la plataforma educativa, el alumnado se conecta a través de sus ordenadores en el centro y desde casa y sus trabajos van siendo enviados al sistema, de manera que todos y todas puedan ver lo que hacen y compartir sus conocimientos y experiencias en colectividad. El alumnado, en sus actividades de aprendizaje musical genera todo tipo de materiales con software de diversa índole en el sistema operativo Guadalinux (de libre distribución). El trabajo se articula más globalmente por la participación del

alumnado en un proyecto de creación de una orquesta escolar, en la que no sólo intervienen ellos sino también músicos profesionales, colaboradores y el propio profesorado. Además de la música elegida e interpretada conjuntamente, en este modelo pedagógico ellos/as crean sus propias coreografías y participan en encuentros musicales con alumnado de otros centros. De esta forma concebimos la acción conjunta y orquestada del alumnado, en la línea de la *teoría de la acción comunicativa* de Habermas (2001), como acción orientada a la comunicación y al entendimiento, a través del intercambio y la colaboración.

Los resultados de este estudio con EDM y su utilidad para la toma de decisiones educativas constituyen el debate principal de cara a la valoración metodológica de estas técnicas. Como resumen básico podemos decir que los modelos estadísticos generados con la aplicación de la minería de datos educativos, reflejan la necesidad de rediseñar la estructura de las plataformas, de forma que ahorremos tiempo y esfuerzo en la exploración y el aprendizaje de la navegación, tratando de hacer el diseño de las plataformas “más transparente” y menos tortuoso para el alumnado; todo ello, con el objetivo de optimizar el aprendizaje, facilitar un mayor aprovechamiento y mejorar la eficacia educativa.

2. Material y método

2.1 Muestra de datos

En nuestra investigación se trabaja con una muestra de alumnado de 250 sujetos-usuarios/as, oficialmente matriculados en el centro de la aplicación y asistentes a clases. Nos encontramos ante una población con diversidad de niveles socioeconómicos de procedencia, que cursa la Educación Secundaria Obligatoria y con edades comprendidas entre 12 -16 años.

La plataforma educativa con la que trabaja el alumnado en nuestro Modelo educativo almacena casi 50.000 registros y alrededor de 100 variables de estudio diferentes a las que sometemos en primer lugar a una limpieza, selección de la información que nos interesa someter a análisis, la aplicación de descriptivos básicos de frecuencias y modelos estadísticos del campo de la minería de datos educativos. Ante una cantidad de información tan grande el software estadístico tradicional se ve limitado y no posee la potencia necesaria para analizar de forma pormenorizada y exhaustiva la información correspondiente a los miles de registros del proceso de uso con los que contamos. Dicha circunstancia nos lleva a plantear la necesidad del uso de un tipo de software específico para este tipo de tareas que nos permita trabajar con comodidad y precisión, de forma fiable y válida con los datos educativos almacenados en las plataformas virtuales de aprendizaje.

2.2 Metodología

Se utiliza una metodología de investigación cuantitativa, haciendo uso de herramientas de recogida de información y análisis del campo del Educational Data Mining. A partir de los datos registrados en tablas en el sistema de gestión del conocimiento, realizamos una limpieza de la información y selección de variables de las que podemos extraer conocimiento útil y significativo de los procesos de aprendizaje online que han tenido lugar. Posteriormente confeccionamos modelos estadísticos que explican dichos procesos de uso.

2.3 Software empleado en el proceso de investigación

En nuestro trabajo con minería de datos educativos empleamos de forma básica tres herramientas informáticas que son SPSS Clementine (versión 11.1), Weka (versión 1.4) y Statistica (versión 8.0). Los dos primeros programas los empleamos para la modelización estadística y el último posee posibilidades de configuración para gráficos tridimensionales que favorecen la visualización de los datos con los que trabajamos.

2.4 La minería de datos educativos en el proceso de evaluación del aprendizaje on-line

En la actualidad en la minería de datos educativos hay dos líneas de trabajo, la primera de ellas es la reingeniería informática de algoritmos para la minería de datos (Romero, Ventura y García, 2008) y la segunda, en la que ubicamos nuestro trabajo, plantea la implementación de técnicas de la minería de datos al campo de la educación y la evaluación de aprendizajes online.

Nuestro estudio plantea la evaluación del proceso de uso de un sistema de gestión del conocimiento basado en las TIC (inspirado en el *Modelo Bordón*) mediante las técnicas de la minería de datos educativos. La minería de datos procede del campo del marketing, la banca y los negocios y es empleada para analizar grandes volúmenes de información registrada en tiempo real en bases de datos. La aplicación de estas técnicas supone la generación de distintos modelos estadísticos descriptivos y predictivos a partir de los datos que quedan registrados y almacenados en esas plataformas en tablas con numerosos registros y variables. En primer lugar, la tarea que realizamos consiste en limpiar y seleccionar las variables significativas para el estudio y confeccionar modelos estadísticos con los datos con los que contamos. A través de esta forma de trabajar con los datos educativos conseguimos darle la importancia y el significado que esa ingente cantidad de información registrada tiene y nos permite extraer conocimiento útil sobre el proceso de uso de una herramienta telemática para la formación musical en red.

2.5 Instrumentación y análisis

En este apartado hacemos una evaluación cuantitativa de la actividad generada en la plataforma educativa-musical. Esta tarea de evaluación sigue un método que denominamos: *Minerización de datos educativos musicales generados en plataformas online*. Para ello, contamos con herramientas de la estadística clásica y la minería de datos educativos. En cuanto a la estadística clásica hacemos uso de análisis descriptivos de frecuencias. La minerización (término tomado de Sierra, 2006) educativa o minería educativa de datos, consiste en la aplicación de modelos estadísticos exploratorios y predictivos orientados a la extracción de patrones de comportamiento significativos de las *tablas que registran toda la actividad* en las plataformas educativas². La minerización abarca múltiples técnicas que van desde la limpieza y selección de variables, la reducción de la dimensionalidad, la segmentación, la clasificación, las reglas de asociación y opciones de visualización de los datos. Para completar este método incluimos procedimientos de minerización complementarios, dependiendo de la naturaleza de los datos obtenidos (categóricos, numerales, ordinales, nominales...).

De todos los métodos con los que contamos en minería de datos educativos, empleamos el Modelo K-Medias para la obtención de conglomerados o grupos de actividad deficiente o anormal (outliers), el modelo basado en gráfico de malla para medir el grado de correlaciones entre variables que reflejan la actividad de la plataforma virtual y la elaboración del modelo A-priori para la obtención de una regla predictiva con los datos del proceso de uso de la herramienta de aprendizaje online. Queremos destacar que el método que seguimos para trabajar con datos educativos, es transportable a cualquier otra materia distinta de la música, o nivel diferente de la Educación Secundaria Obligatoria, por lo que la utilidad de nuestro método radica en su compatibilidad y portabilidad, siempre y cuando empleemos sistemas de gestión del conocimiento de la naturaleza del nuestro (como Blackboard, Helvia, WebCT...).

3. Resultados de la investigación

3.1 Análisis estadísticos descriptivos básicos.

En cuanto a la *actividad en la plataforma del profesor-investigador y el alumnado*, el 20,2 por ciento (9.934 registros) de la actividad de la plataforma es realizada por el profesor y el 79,8 por ciento (39.161 registros) por el total del alumnado, es decir, una quinta parte,

² El término KDD significa en inglés Knowledge Databases Discovery (extracción de conocimiento [útil] a partir de bases de datos).

aproximadamente de la actividad global. A continuación mostramos estos datos en una tabla y un gráfico de barras. El número total de registros de las tablas de la plataforma educativa es de 49.095, contando con la actividad del profesor y del alumnado.

| | | Frecuencia | Porcentaje | Porcentaje Válido | Porcentaje acumulado |
|---------|----------|------------|------------|-------------------|----------------------|
| Validos | Profesor | 9934 | 20,2 | 20,2 | 20,2 |
| | Alumnado | 39161 | 79,8 | 79,8 | 100,0 |
| | Total | 49095 | 100,0 | 100,0 | |

Tabla 1. Frecuencia de participación en la plataforma del profesor y el alumnado.

En relación a la *actividad en la plataforma por cursos (profesor-investigador y alumnado)*, en primero el porcentaje participación del 20,8 por ciento (10.182 registros), en segundo el 27,9 por ciento (13.362), en tercero el 38,6 por ciento (18.914) y en cuarto el 12,8 (6.270 registros). A continuación mostramos estos datos en una tabla.

| | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|---------|-------|------------|------------|-------------------|----------------------|
| Validos | 1º | 10182 | 20,8 | 20,8 | 20,8 |
| | 2º | 13662 | 27,9 | 27,9 | 48,6 |
| | 3º | 18914 | 38,6 | 38,6 | 87,2 |
| | 4º | 6270 | 12,8 | 12,8 | 100,0 |
| | Total | 49028 | 100,0 | 100,0 | |

Tabla 2. Actividad de la plataforma educativo-musical distribuida por cursos.

Tal y como hemos comprobado en los análisis descriptivos, el profesor acumula la quinta parte de la actividad registrada en la plataforma y el curso con mayor actividad ha sido tercero de la ESO.

3.2 Limpieza de la información: reducción y selección de variables

Tras la tarea de exploración de las tablas de la plataforma educativa, concluimos que aquellas variables que nos arrojan una información más valiosa en nuestra investigación son: *userid* (código numérico que identifica al alumnado participante en la plataforma educativa), *course* (curso del usuario/a, primero, segundo, tercero o cuarto de ESO), *module* (módulo de la plataforma utilizado: foro, chat, lección, cuestionario...), *action* (acción realizada dentro de cada módulo de la plataforma: ver, subir archivo, descargar...), *post* (número de mensajes enviados a los foros), *quiz* (realizar test de conocimientos musicales), *grade* (puntuación obtenida por los usuarios/as en las distintos módulos de la plataforma: foros, lecciones o cuestionarios). A continuación mostramos en un mapa conceptual las variables de estudio de la plataforma educativa-musical:

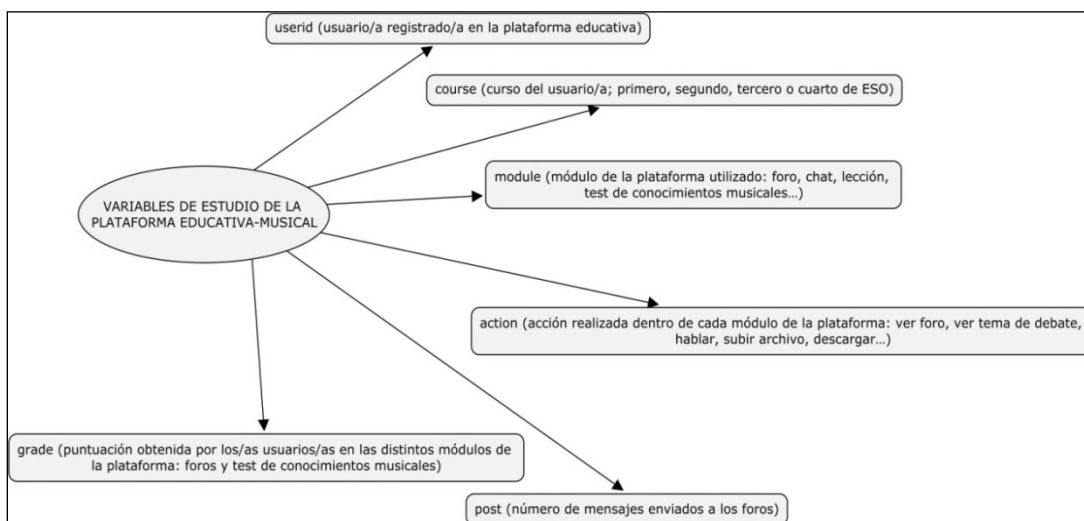


Figura 1. Variables de estudio de la plataforma educativa-musical.

De todas las variables que consideramos significativas o relevantes, mostramos en este artículo aquellas que nos sirven para la modelización estadística que presentamos a continuación.

3.3 Generación de modelos estadísticos con la minería de datos educativos

Las tareas de minerización musical que planteamos, consisten en la extracción de patrones de comportamiento de la información registrada en la base de datos de nuestra plataforma educativa. Dichas tareas, se basan en aprehender conocimiento útil y significativo (partiendo de los datos almacenados en las tablas de registro de actividad de nuestro soporte informático), mediante la aplicación de modelos estadísticos variados. En nuestro trabajo, las posibilidades de aplicación del *Data mining* en la evaluación de aprendizajes online son muy variadas ya que abarcan múltiples técnicas exploratorias, gráficas y clasificatorias. De todos los modelos que hemos hallado a lo largo de nuestra investigación, mostramos a continuación *aquéllos que reflejan una información más relevante y significativa sobre el comportamiento de los datos en el proceso de uso de la plataforma virtual de aprendizaje: modelo K-Medias, Gráficos de malla y modelo predictivo A-priori*. En cuanto al tiempo invertido en la generación de estos modelos, en ningún caso supera los 5 segundos, lo cual nos posibilita una gran cantidad de combinaciones y variaciones en el análisis hasta alcanzar un modelo óptimo que ofrezca la máxima validez y fiabilidad.

- a) *Modelo K-Medias del registro general de datos la plataforma educativa (grupos de actividad según frecuencia de uso de la plataforma educativa)*
- b)

El objetivo del *modelo estadístico K-Medias* es extraer información relevante acerca de los conglomerados o clústeres que agrupan en distintas categorías aquellos datos con un comportamiento parecido o similar. En nuestro caso, hacemos un estudio de la frecuencia de uso de la plataforma educativa a partir del registro general de actividad del sistema.

Los grupos de actividad o conglomerados que detectamos en la plataforma son 10: el *conglomerado 1* con 6.766 registros; *conglomerado 2* con 1.659 registros; *conglomerado 3* con 1.939 registros; *conglomerado 4* y *conglomerado 8* con tan sólo 159 registros y 133 registros respectivamente, siendo los conglomerados menos numerosos en cuanto a la cantidad de registros muestran grupos de actividad deficiente; *conglomerado 5* con 2.704 registros; *conglomerado 7* con 1.343 registros; *conglomerado 9* con 3.612 registros; *conglomerado 10* con 4.568 registros. Seguidamente, mostramos en un gráfico de barras el número de registros de cada uno de los conglomerados hallados:

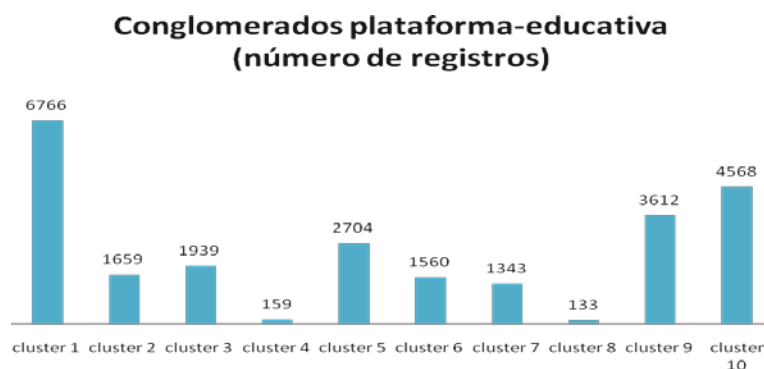


Figura 2. Conglomerados de la plataforma educativa (según el número de registros acumulados).

Tras la aplicación de esta tarea de minerización descubrimos que los grupos 1, 10 y 9 representan una alta frecuencia de uso de la plataforma virtual de aprendizaje, los grupos 2, 3, 5, 6 y 7 una frecuencia de uso normal y los grupos 4 y 8 una baja frecuencia de uso o actividad deficiente de la plataforma (outliers), o sea, un comportamiento anómalo o distinto del resto de los datos.

b) Modelo de malla direccional del registro general de actividad de la plataforma virtual (correlaciones entre variables)

Los *nodos de malla direccional* se utilizan para ilustrar la fuerza de las relaciones existentes entre los valores de dos o más campos o variables. Las conexiones se muestran en un gráfico con distintos tipos de líneas para indicar conexiones de creciente fuerza o grado de correlación. En nuestro caso medimos el grado de correlación entre dos variables: *módulo* y *acción*. Las líneas más gruesas indican un índice mayor de relación y las más débiles un menor índice de relación.

En los enlaces fuertes observamos que la acción “ver curso” destaca por encima de las demás con 8.663 enlaces, seguida por “ver foro de discusión”, con 7.433 enlaces, “ver foro”, con 5.310 enlaces y “hablar en el chat” con 4.221 enlaces.

Los enlaces medios reflejan que la acción “ver curso” tiene 34 enlaces, la acción “editar sección del curso” 24 enlaces, “actualizar recurso” 24 enlaces, “ver gráfico de la encuesta” 23 enlaces, “ver informe de la encuesta” 24 enlaces.

En los enlaces débiles, la información vertida, hace referencia a tareas de mantenimiento de la herramienta de aprendizaje.

El análisis del proceso de uso de la plataforma, a través del estudio de los enlaces entre variables muestran *acciones exploratorias de la herramienta* (ver un recurso o explorar una sección de la plataforma educativa), *acciones de aprendizaje* (hablar en los chats y enviar trabajos a los foros de comunicación) y de *gestión o mantenimiento del sistema* (optimizar o actualizar un recurso del sistema).

Seguidamente mostramos el gráfico de malla direccional con diseño circular que muestra el enlace fuerte entre el módulo *course* y la acción *view*, que refleja el alto grado de correlación entre estas dos variables, que denotan la importancia de las acciones exploratorias en la plataforma virtual de aprendizaje:

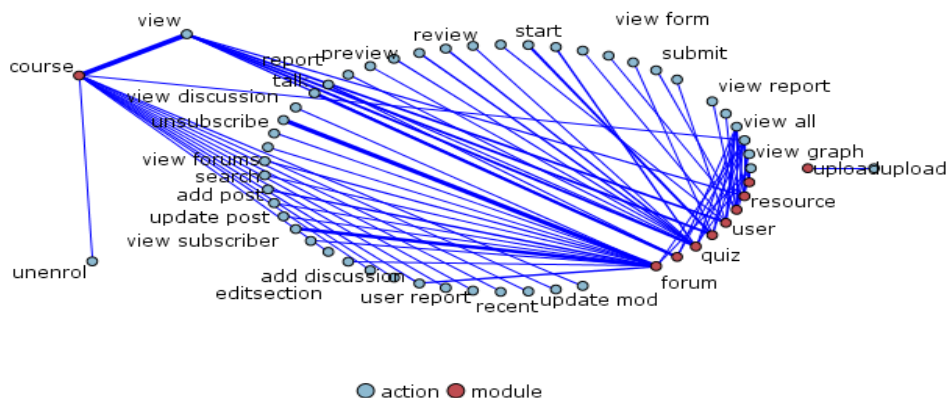


Figura 3. Enlace fuerte entre el módulo *course* y la acción *view* (gráfico tomado de Espigares y García, 2010).

c) *APRIORI del registro general de actividad de la plataforma educativa (elaboración de un regla predictiva)*

El objetivo de la aplicación del Modelo estadístico *A-priori* es extraer un conjunto de reglas de los datos y destaca aquellas reglas con un mayor contenido de información sobre el proceso de uso de la plataforma educativa. A priori ofrece cinco métodos diferentes para la selección de reglas y utiliza un sofisticado esquema de indización para procesar eficientemente grandes conjuntos de datos. En los problemas de mucho volumen, A-priori se entrena más rápidamente que GRI, no tiene un límite arbitrario para el número de reglas que puede retener y puede gestionar reglas que tengan hasta 32 precondiciones. A-priori requiere que todos los campos de entrada y salida sean categóricos, pero ofrece un mejor rendimiento ya que está optimizado para este tipo de datos.

El resultado que obtenemos tras la aplicación del Modelo estadístico A-priori es que hallamos una única regla que nos indica que “si el antecedente es el módulo *userid*, con un 19,047% de grado de soporte (del total de registros acumulados en la plataforma) y un 90,228% de grado de confianza el consecuente es la acción *view*”.

| <i>Consecuente</i> | <i>Antecedente</i> | <i>% de soporte</i> | <i>% de confianza</i> |
|--------------------|--------------------|---------------------|-----------------------|
| Action=view | Module=course | 19,047 | 90,228 |

Tabla 3. Modelo APRIORI.

El resultado del análisis para la obtención de dicha regla nos arroja la siguiente información: contamos con 16.118 transacciones válidas, un soporte mínimo y máximo del 19,047%, confianza máxima y mínima del 90,228%.

4. Discusión

Las conclusiones que obtenemos a partir del estudio realizado con técnicas y herramientas de la minería de datos educativos aplicadas al Moodle para el aprendizaje musical online, las recogemos a continuación en los siguientes puntos:

1) Las *técnicas de muestreo estadístico aplicadas a los cerca de 50.000 registros* del proceso de uso del sistema de gestión del conocimiento que estudiamos, reflejan que ordenados de mayor a menor según la *frecuencia de uso de la plataforma por cursos* son: tercero de la ESO, segundo de la ESO, primero de la ESO y cuarto de la ESO. En esta información, *detectamos anomalías en determinados grupos que no han realizado correctamente las tareas*, lo cual nos

permite monitorizar los objetos de aprendizaje en los que su uso y rendimiento académico, por parte del alumnado, presenta un comportamiento anómalo o deficiente.

2) En cuanto a los *niveles de uso de la plataforma educativa por parte del profesor y del alumnado*, tras la aplicación del muestreo estadístico con el total de los registros del proceso de uso de Moodle, observamos que el profesor-investigador de la asignatura desempeña aproximadamente una quinta parte de la actividad registrada en la plataforma (un 20% más o menos), quedando las otras cuatro quintas partes registradas por el alumnado. Un mayor uso, desde las casas incrementaría notablemente esta cierta desproporción entre la actividad del profesor en la plataforma y el conjunto del alumnado de primero a cuarto de ESO. También se hace patente la escasa carga horaria de la asignatura de música en la Educación Secundaria y la necesidad de una mayor carga lectiva semanal. En ese sentido y para potenciar el uso de nuestra herramienta informática, realizamos durante el curso dos tutorías virtuales con el alumnado para orientarlo en su uso y aprendizaje con la misma, teniendo lugar una al final del segundo trimestre del curso y otra al final del tercero a través de la herramienta de aprendizaje “chat de intercambio de la información” que posee el sistema.

3) Los *modelos estadísticos* empleados, así como las variables de estudio, seleccionadas por sus alto valor y relevancia educativa en nuestro estudio, de *conglomerado (K-medias)* y *predictivo (Apriori predictivo)* nos ofrecen información complementaria a las técnicas de muestreo y recuento de los registros del proceso de uso. El modelo tipo clúster, K-medias nos permite conocer grupos de inactividad y aprendizaje deficiente, así como anomalías en la adquisición de conocimiento. En nuestro estudio, de los 10 conglomerados que obtenemos a partir del estudio de la frecuencia de uso de la plataforma, los grupos de actividad 4 y 8 reflejan un comportamiento anómalo o distinto del resto de los datos (outliers). A partir de esta información, podemos monitorizar a estos grupos de actividad, prestarle una atención especial, desarrollar nuevas actividad de enseñanza-aprendizaje y optimizar el funcionamiento del sistema. El modelo A priori predictivo muestra que a través del análisis del cómputo global de los registros del proceso de uso de la plataforma educativa, obtenemos un regla predictiva que establece que con un grado de más de un 90 por ciento de confianza, los usuarios tienden a realizar acciones de tipo exploratorio, como ver o navegar por el sistema.

4) En cuanto a la *monitorización de los módulos de aprendizaje* de la plataforma, prestamos especial atención a los *foros para la Educación Musical*, por registrar la mayor parte de la actividad. El módulo que presenta una actividad más intensa es el “foro de intercambio de la información para el alumnado”. Esto refleja que el alumnado ha trabajado de forma correcta, en

su mayoría, empleando esta herramienta de comunicación, enviando sus mensajes y documentos para que éstos fueran evaluados por toda la comunidad de aprendizaje musical online. Dentro de este objeto digital de aprendizaje, las acciones que presentan una mayor frecuencia de uso son las de tipo exploratorio, como *ver foros* o *ver temas de debate de los foros*.

5) Las *acciones de tipo exploratorio* apreciadas en la plataforma educativa, constituyen un nexo de unión entre el pseudoaprendizaje, los preconceptos y la adquisición de auténticos aprendizajes, es decir, la indagación y la exploración como paso previo a la obtención de metaconocimiento musical o aprendizajes musicales de alto nivel. En este sentido, nuestro estudio detecta la necesidad de un *mayor tiempo de trabajo con los sistemas LMS para la Educación Musical* y la necesidad de *desarrollar, mejorar y optimizar el empleo y los hábitos de trabajo con estas tecnologías del conocimiento conectadas directamente con la filosofía TIC 2.0*, que promueve el empleo de este tipo de herramientas de aprendizaje online en todos los ámbitos del conocimiento.

6) El empleo de todas las técnicas, algoritmos y modelos estadísticos de la *minería de datos educativos que empleamos en nuestro estudio nos han permitido trabajar y analizar la información del proceso de uso de la plataforma educativa en todo momento, de forma rápida y fiable* con una gran cantidad de información educativa para cuyo análisis, el EDM se presenta como una herramienta de trabajo adaptada y adecuada a las necesidades del tipo y la cantidad de información que se registran actualmente en los sistemas del conocimiento. El trabajo con grandes volúmenes de información, con numerosas variables y registros para analizar plantea el *uso inevitable de otro tipo de software más adecuado y distinto del software clásico y convencional, por ser más rápido y flexible, básicamente por su arquitectura, solidez y técnicas de muestro de la información*. Éstas mejoran la eficiencia y el rendimiento de populares programas y aplicaciones en el campo educativo y de las ciencias sociales, muy utilizados pero en la actualidad insuficientes para este tipo de tareas de extracción de conocimiento útil y significativo a partir de los registros de uso de sistemas de gestión del aprendizaje.

En resumen, en la evaluación que hacemos del proceso de uso de la plataforma virtual de aprendizaje musical, contamos con *diferentes herramientas de análisis que nos posibilitan medir la actividad mediada por las nuevas tecnologías*, que potencian el entendimiento y configuran un modelo de sociedad, en el que la comunicación, mediante el uso de plataformas de teleformación, se potencia y se enriquece a través del aprendizaje colectivo, que propicia la creación de redes sociales virtuales para el aprendizaje musical en escuelas de Educación Secundaria Obligatoria y

que a través de la minería de datos educativos es posible estudiar de manera exhaustiva, rigurosa y detallada.

Por todo ello, planteamos la *necesidad de que en el futuro se integren las herramientas de análisis estadístico de la minería de datos educativos dentro de las propias plataformas de teleformación*, para que en cualquier momento dispongamos de la tecnología necesaria para detectar deficiencias en el proceso de enseñanza-aprendizaje online. De esta forma se podrán monitorizar los objetos digitales de aprendizaje musical, optimizar la configuración de la navegación en el sistema de teleaprendizaje y procurar que la adquisición de conocimiento útil en la asignatura mediante nuestro modelo educativo sea más fácil, sencilla y directa. Esto proporcionará que *no haya necesidad de tener que emplear herramientas externas a la propia plataforma y hacer un uso óptimo de la misma para que el alumnado reciba un feedback inmediato de su proceso de aprendizaje* de cara a la obtención de mayores destrezas y competencias musicales.

Referencias bibliográficas

- Area Moreira, M. (2004). *Los medios y las tecnologías en la educación*. Madrid: Pirámide.
- Chien-Ming Chen, Cheng-Hao Ma, Bin-Shyan Jong, Yen-Teh Hsia y Tsong-Wuu Lin (2008). Using Data Mining to Discover the Correlation between Web Learning Portfolios and Achievements. *Proceedings of the 38th ASEE/IEEE Frontiers in Education Conference* (Saratoga Springs, Nueva York).
- Cole, M. (1999). *Psicología Cultural*. Madrid: Morata.
- Cox, R. (2009). Data mining PATSy: Insights into health science students' clinical reasoning. *Proceedings of the 13th. Biennial Conference* (Universiteit Van Amsterdam).
- Daniels, H. (2003). *Vygotski y la pedagogía*. Barcelona: Paidós.
- De Aguilera, Miguel y Adell, J. A. (2010). Mediaciones en el nuevo escenario digital. *Comunicar*, 34, (17), 10-14.

De Pablos, J. (1996). *Tecnología y Educación. Una aproximación sociocultural*. Barcelona: Cedecs.

Espigares Pinazo, M. J. y García Pérez, R. (2006). Educación musical con TICs en Escuelas Multiculturales. Actas del *I Jornadas Internacionales de Educación Intercultural (VI Jornadas de Educación Intercultural “convivencia y mediación intercultural”)* (Universidad de Almería).

Espigares Pinazo, M. J. y García Pérez, R. (2010). Evaluación de un Modelo de gestión del conocimiento educativo-musical: El Modelo Bordón. *LEEME. Lista Electrónica Europea de Música en la Educación*, 25, 49-64.

Habermas, J. (2001): *Teoría de la Acción Comunicativa*. Vol. II. Madrid: Taurus.

Romero, C., Ventura, S. y García E. (2008). Data Mining in Course Management Systems: MOODLE Case Study and Tutorial. *Computers & Education*, 51 (1), 368-384.

Sierra Araujo, B. (2006). *Aprendizaje automático: conceptos básicos y avanzados. Aspectos prácticos utilizando el software Weka*. Madrid: Pearson Prentice Hall.