

APLICACIÓN DE MÉTODOS DE DISEÑO CENTRADO EN EL USUARIO Y MINERÍA DE DATOS PARA DEFINIR RECOMENDACIONES QUE PROMUEVAN EL USO DEL FORO EN UNA EXPERIENCIA VIRTUAL DE APRENDIZAJE

(APPLICATION OF USER-CENTERED DESIGN METHODS AND DATA MINING TO DEFINE RECOMMENDATIONS, WHICH FOSTERS THE USE OF THE FORUM IN A VIRTUAL LEARNING EXPERIENCE)

Priscila M. Valdiviezo
Unidad de Virtualización, UTPL (Ecuador)

Olga C. Santos
Jesús G. Boticario
Grupo de Investigación aDeNu, Departamento de Inteligencia Artificial, ETSI Informática, UNED (España)

RESUMEN

La adopción de sistemas recomendadores en ambientes virtuales de aprendizaje se está convirtiendo en una alternativa; para lograr la adaptación automática requerida, para atender las necesidades de aprendizaje de los estudiantes. Con los datos de interacción, que proveen estos ambientes es posible encontrar indicadores que con la aplicación de técnicas de minería de datos y aprendizaje automático se pueda identificar información relevante, para la definición de recomendaciones. En esta investigación, hemos aplicado técnicas de aprendizaje no supervisado, para la identificación de patrones comunes de interacción con los foros disponibles en un curso de la plataforma OpenACS/dotLRN. Esto facilitará la definición de recomendaciones que ayuden a mejorar la experiencia de aprendizaje de los estudiantes.

Palabras clave: sistemas adaptativos de educación, OpenACS/dotLRN, modelado de usuario, minería de datos y aprendizaje automático, técnicas de agrupamiento automático (del inglés “clustering”), recomendaciones.

ABSTRACT

Recommender systems in learning virtual environments are increasingly becoming a feasible approach to provide the adaptive support required to attend students' learning needs. With interaction

data obtained from these virtual environments, it is possible to find indicators where data mining and machine learning techniques can be applied to identify relevant information that allows for the definition of recommenders. In this research we have applied unsupervised learning techniques to identify common interaction patterns with available forums in the OpenACS/dotLRN platform course. In this way, it will allow to define recommendations which help to improve the learning experience of students.

Keywords: adaptive systems in education, OpenACS/dotLRN, user modelling, data mining and machine learning techniques, clustering techniques, recommendations.

La aparición de los sistemas recomendadores, se debe principalmente a la necesidad de reducir la gran cantidad de información, que el usuario tiene que procesar a fin de encontrar algo interesante (Adomavicius & Tuzhilin, 2005). Uno de los campos de aplicación, en los cuales estos sistemas pueden ser de utilidad es el e-learning, en donde en los últimos años se están llevando a cabo avances significativos. Hoy en día los usuarios de esta modalidad de estudio, (e-learning) tienen una gran variedad de capacidades, intereses, nivel de experiencia en el uso de las Tecnologías de la Información y la Comunicación (TIC). No obstante, considerando la heterogeneidad de los usuarios de esta modalidad no basta sólo con los servicios que se brinden para hacer posible el proceso de aprendizaje; es recomendable también dar respuestas adaptadas a cada usuario, de tal forma que se sienta guiado en el proceso de aprendizaje. Una forma de lograr esta adaptación es, mediante el uso de los llamados sistemas recomendadores, los cuales pueden ser entendidos como “sistemas basados en tecnología que ofrecen recomendaciones personalizadas a los usuarios, las mismas que pueden ser generadas a partir de las opiniones y acciones de otros usuarios con gustos similares” (Sanjog & Mahanti, 2008).

Algunos trabajos en escenarios de e-learning proponen estrategias de recomendación, con el propósito de sugerir al alumno la realización de actividades de aprendizaje, el seguimiento de rutas de navegación, la consulta de material educativo, o cualquier otro recurso que permita potenciar y mejorar su proceso de aprendizaje (Drachler et al., 2008; Santos & Boticario, 2008). En base a esto, los estudiantes pueden recibir ayuda en el desempeño de las tareas del curso, con el propósito de: a) evitar bloqueos, b) mejorar el desempeño de los procesos de aprendizaje facilitando los contenidos del curso más apropiados y las rutas de aprendizaje adaptadas a las necesidades de los estudiantes y c) promoviendo la colaboración entre pares (Santos & Boticario, 2008). Todo esto en orden a mejorar su aprendizaje de forma eficiente.

En concreto, conviene enfatizar las ventajas de promover al aprendizaje activo. Schank (1995) defiende que “la mejor forma de aprender es haciendo”. De esta forma,

el alumno aprende cuando construye, aprende a pensar, a decidir por sí mismo y expresar lo que uno ha aprendido de un modo persuasivo”. De esta forma, además de proporcionar conocimiento al alumno mientras aprende a su ritmo y según sus intereses; se le ha de proporcionar al estudiante los recursos necesarios, para poder no sólo resolver la tarea y elaborar la solución, sino también poner en común la solución y permitir el debate colectivo, para reconstruir entre todos la solución.

En este sentido, los sistemas recomendadores se convierten en una alternativa, para brindar una adaptación personalizada a los estudiantes. En este contexto educativo el sistema recomendador no sólo brinda soporte al estudiante durante la ejecución de un curso, sino que también ayuda a reducir la carga de trabajo del profesor, dado que permite complementar la actuación de éste ante cada alumno y situación dada. Además, las recomendaciones que se brindan al estudiante, están basadas en sus preferencias, conocimientos, intereses, información histórica de otros alumnos de características similares, etc., es decir, que estos sistemas toman como entrada “el perfil del usuario (el cuál puede ser construido dinámicamente desde las interacciones del usuario) y el contexto actual (ej. curso, objetivos, herramientas de la plataforma...)”, para brindar recomendaciones apropiadas al estudiante.

Para llevar a cabo la oferta de recomendaciones, Santos et al. (2009) proponen métodos de diseño centrado en el usuario para conocer en qué situaciones conviene ofrecer dichas recomendaciones. Ante esto, podemos decir que las recomendaciones están sujetas a ciertas condiciones que deben ser identificadas previamente, ya sea mediante el análisis de los datos que se tenga del usuario y/o a través de diferentes técnicas de aprendizaje automático, que permiten un proceso automatizado de análisis, y que además permitan deducir los valores de ciertas características útiles para la condición. En esta investigación se presenta una propuesta basada en este enfoque, que consiste en el uso de un sistema recomendador semántico (Santos & Boticario, 2008), basado en técnicas de aprendizaje automático, junto con el análisis de las interacciones que realizan los usuarios durante su participación en cursos online, con el objetivo de mejorar la experiencia de aprendizaje de los estudiantes.

En este artículo nos centraremos en aplicar dicho planteamiento para definir recomendaciones centradas en el uso de los foros en un contexto real de uso, que involucra a estudiantes de la carrera de Ingeniería en Sistemas Informáticos y Computación de la Universidad Técnica Particular de Loja en Ecuador; y que combina la recogida de información tanto de forma directa del usuario (encuestas y entrevistas) como indirecta (observaciones y análisis de las trazas generadas). Esta información se analiza tanto a nivel estadístico como aplicando técnicas de

minería de datos, con el fin de proponer recomendaciones relativas al uso del foro que ayuden a mejorar la experiencia de aprendizaje, y que se presenta como un caso de estudio para promover el uso de la herramienta foro que aplica la metodología de diseño de recomendaciones para escenarios educativos, propuesto por el grupo de investigación aDeNu1 de la UNED.

Este artículo se inicia con una revisión del estado del arte sobre los sistemas recomendadores en el campo educativo y las técnicas de aprendizaje automático que pueden ser utilizadas para la clasificación no supervisada dentro del contexto del elearning; seguidamente, se presenta el contexto y planteamiento de la experiencia donde se describe el proceso para la obtención de indicadores de colaboración para la aplicación de técnicas de agrupamiento automático (generalmente conocida por su designación en inglés “clustering”); así mismo, se presentan los resultados de la experiencia donde se hace una comparación y discusión de los mismos, un análisis crítico abordando posibilidades de mejora para futuras experiencias. Finalmente, se presentan las conclusiones y trabajos futuros.

Estado del arte

En un entorno e-learning el sistema recomendador debe tratar con información del estudiante, como por ejemplo: información personal (género, edad, idioma, etc.), características de accesibilidad, objetivos de aprendizaje, preferencias, etc., la cual se encuentra reflejada en un modelo de usuario, que se constituye en un elemento importante en estos sistemas, ya que las recomendaciones se generan a partir de la información que se encuentra almacenada en el mismo.

Por consiguiente, contar con un sistema recomendador en entornos educativos, implica entender el contexto individual del estudiante (o grupos de estudiantes) y las condiciones y reglas del dominio (Drachler et al., 2008).

Sistemas recomendadores en educación

La línea de investigación de los sistemas recomendadores ha dado origen a diferentes proyectos dentro de la educación superior, entre estos podemos citar a EU4ALL² dirigido a estudiantes con discapacidad en el ámbito de estudios superiores, en el cual se pretende cubrir distintos tipos de escenarios educativos susceptibles de adaptación. Este proyecto contempla el desarrollo de un sistema recomendador como una forma de brindar apoyo adaptativo al proceso de aprendizaje.

Los sistemas recomendadores requieren de técnicas para seleccionar ítems, opciones, servicios, etc. que mejor se ajustan a las necesidades o preferencias del usuario y en base a la información de la que disponen estos sistemas se puede decidir qué técnicas seleccionar y para qué situaciones.

De acuerdo con Drachsler et al. (2008), entre las técnicas de recomendaciones se pueden citar:

Basadas en Modelo. Los sistemas recomendadores basados en estos modelos requieren de una gran cantidad de datos para estimar sus modelos y así proporcionar una recomendación exacta.

Basada en Memoria. Analizan todos los datos de usuarios o ítems para calcular las recomendaciones, éstas pueden ser clasificadas como:

- **De filtrado colaborativo:** que permiten sugerir a un estudiante en particular, un ítem o una acción, en base a la opinión de otros estudiantes con perfiles similares. Esta técnica usa el desempeño colectivo de todos los estudiantes de la red de aprendizaje para identificar actividades de aprendizaje de apropiadas en un contexto dado y permite a los estudiantes beneficiarse de las experiencias de otros estudiantes. Además, puede estar basada en reglas pedagógicas que formen parte de la estrategia de recomendación, así las características de los estudiantes podrían ser tomadas en cuenta para asignar a los estudiantes en grupos (Drachsler et al., 2008).
- **Basada en contenido:** con esta técnica, las recomendaciones se realizan en base al perfil del estudiante y a las evaluaciones que el mismo realice de un curso, material educativo, o cualquier otro recurso de aprendizaje.
- **Enfoques híbridos:** combinan las características de ambos sistemas mitigando las limitaciones de cada uno. De acuerdo a las técnicas mencionadas, las recomendaciones se generan ya sea a partir de las opiniones dadas por otros usuarios o bien a partir de las preferencias del usuario objeto de la recomendación.

Drachsler et al. (2008) proponen una combinación de las técnicas de recomendación basada en memoria, como las más convenientes para realizar recomendaciones personalizadas de las actividades de aprendizaje en el contexto de e-learning.

Así mismo, los sistemas recomendadores en el contexto educativo pueden basar sus recomendaciones en la identificación de tipos de estudiantes con la ayuda de técnicas de aprendizaje automático y así proporcionar recomendaciones apropiadas a las características individuales de los estudiantes o las presentadas por grupos de estudiantes (Vialardi et al., 2009; Anaya & Boticario, 2010).

Por otro lado, algunos sistemas existentes optan por las redes semánticas como técnicas de representación de los perfiles de usuario, y cómo estrategias de representación del conocimiento durante el proceso de recomendación, ya que permiten dotar de significado a la información utilizada para este proceso (Yu et al., 2007).

La incorporación de la semántica en los sistemas recomendadores permitiría contar con ciertos procesos de razonamiento capaces de descubrir relaciones importantes entre los ítems disponibles y las preferencias personales de los usuarios. Mediante los métodos de razonamiento semántico, el sistema recomendador no sólo dispone de más conocimiento para poder elaborar sus sugerencias, sino que además, puede combatir de forma efectiva las limitaciones identificadas en las técnicas de recomendación utilizadas de forma genérica en los sistemas comerciales en la Web.

Técnicas de aprendizaje automático y su aplicación en el dominio educativo

En los sistemas educativos las técnicas de aprendizaje automático son una alternativa para clasificar y predecir acciones futuras de los estudiantes en el sistema. En algunos enfoques de sistemas recomendadores, durante el proceso de recomendación se opta por modelar las preferencias de los usuarios mediante técnicas de aprendizaje automático, tales como: redes neuronales, árboles de decisión, redes bayesianas, etc. Estos algoritmos de clasificación supervisada se utilizan comúnmente como parte de las técnicas de minería de datos justamente cuando se cuenta con una enorme cantidad de datos que necesitan ser clasificados y analizados.

Según Anaya y Boticario (2010), cuando el volumen de datos es suficiente, la clasificación de instancias se puede hacer aplicando técnicas de aprendizaje no supervisado, como clustering. De acuerdo a Perera et al. (2007), el clustering permite utilizar múltiples atributos para identificar grupos similares de una manera no supervisada, sin la necesidad de etiquetar los grupos. Por ello, al análisis de clustering se lo conoce como método de clasificación automática no supervisada que

“consiste en encontrar la partición más adecuada del conjunto de entrada a partir de similitudes entre sus ejemplos” (García y Gómez, 2009). Es decir que, la idea de formar clústeres es agrupar elementos en conjuntos homogéneos en función de algunas semejanzas entre ellos y diferentes a los que pertenecen a otros grupos. Una vez formados los clúster se realiza el reconocimiento y descubrimiento de patrones de semejanza, que permiten a la vez llevar a cabo acciones sobre los datos.

Dentro de la técnica de clustering se puede encontrar los siguientes métodos:

De Partición. Estos métodos “dan lugar a una partición mediante la minimización o maximización de algún criterio” (Sierra, 2006, p. 286). La partición obtenida depende del algoritmo y del criterio de optimización que se utilice. Los métodos de partición utilizan directamente una matriz de datos y para poder aplicarlos necesitan que el usuario fije de antemano el número de clústeres o grupos que tendrá la partición.

Entre los métodos de partición están el algoritmo K-means que sigue un procedimiento simple de clasificación de un conjunto de objetos en un determinado número de clústeres determinado a priori (García y Gómez, 2009). K-means representa cada uno de los clústeres por la media de sus puntos, es decir, por su centroide, y así cada clúster es caracterizado por éste, el cual se encuentra en el centro de los elementos que componen el clúster.

Jerárquicos. Como se describe en Sierra (2006, p. 268), estos métodos establecen una jerarquía entre los clúster, es decir, se tiene una sucesión de particiones donde cada partición se obtiene uniendo o dividiendo clústeres. Existen dos tipos de métodos jerárquicos: *de aglomeración*, donde los nuevos clústeres se crean uniendo clúster, y *de división*, parten de un conjunto de elementos y se van dividiendo sucesivamente hasta llegar a elementos individuales, es decir los nuevos clústeres se crean dividiendo clústeres.

Clustering basado en probabilidades, en este caso los objetos tienen cierta probabilidad de pertenecer a un grupo o clúster (García & Gómez, 2009). Un algoritmo de clustering probabilístico es el algoritmo EM (*Expectation Maximization*), que presenta una técnica iterativa del cálculo de máxima verosimilitud que se utiliza para encontrar una estimación del conjunto de parámetros de problemas en los que existen ciertos datos ocultos (Sierra 2006, p. 150). Puede aplicarse en situaciones en las que se desea estimar parámetros desconocidos, y posteriormente intentar maximizar la probabilidad de que cada objeto o elemento pertenezca a un grupo o clúster.

Estas técnicas se han aplicado con éxito en el campo de la educación. Boticario y Anaya (2009), presentan la aplicación del algoritmo EM para construir los clústeres, el cual usa indicadores estadísticos de la interacción de los estudiantes en los foros y técnicas de minería de datos para la obtención de grupos de estudiantes de acuerdo a su nivel de colaboración. Con este algoritmo la tarea de agrupación puede ser vista como un problema de estimación de máxima verosimilitud (Anaya & Boticario, 2010).

Un aspecto importante es que el clustering basado en probabilidades (Talavera & Gaudioso, 2004), es una aproximación que ha ganado popularidad en la literatura trabajando con ambos tipos de datos continuos y discretos. Existe una variedad de algoritmos de clustering, el seleccionar uno u otro dependerá de la aplicación en particular, y del tipo de datos a utilizar.

Para esta investigación se requerirá un algoritmo capaz de trabajar con datos discretos, por ello se probarán los dos tipos de algoritmos mencionados K-means y EM, y se seleccionará el modelo que mejores resultados presente.

Contexto de la experiencia

Considerando el trabajo presentado por Santos y Boticario (2008), diferentes tipos de recomendaciones pueden ser definidas de acuerdo a las funcionalidades disponibles en un sistema de gestión del aprendizaje (learning management system – LMS). De acuerdo a dichos tipos, se presentan una serie de recomendaciones que pueden ser dadas al usuario, considerando para su aplicación los elementos definidos en el modelo de recomendaciones: las categorías, las técnicas, el origen, la explicación, las restricciones y las condiciones.

Dichos elementos permiten clasificar una recomendación según su categoría, que puede ser generada por una técnica (o combinación de técnicas). La oferta de cada una de las recomendaciones está sujeta a ciertas condiciones y restricciones. Las condiciones hacen referencia a: preferencias de accesibilidad, acciones sobre objetos, interacción con los datos, nivel de conocimiento, nivel de colaboración y de tecnología; además se añade una explicación del por qué de la recomendación. La categoría, la técnica, la explicación, así como el origen y la razón forman la justificación de la recomendación.

Para la definición de estas recomendaciones el sistema necesita de información del estudiante acerca de su interacción con el LMS y del modelo del usuario, para

de esta manera recomendar información cercana a sus necesidades. La oferta de recomendaciones se realiza usando información relacionada a ciertos indicadores que pueden ser obtenidos directa o indirectamente de las interacciones de los usuarios (Santos & Boticario, 2004). Estos a la vez pueden ser recogidos de los datos almacenados en los archivos de log del servidor con el objeto de identificar atributos que luego puedan ser incorporados dentro del perfil del usuario y que puedan ser utilizados para crear una serie de recomendaciones (Quincey et al., 2009).

Estos trabajos previos son los que han alimentado la experiencia realizada en esta investigación y toman como base la necesidad detectada de recomendar a los estudiantes que visiten y contribuyan en el foro, que fue la principal herramienta de comunicación en la experiencia llevada a cabo en esta investigación. Como se señala en Anaya y Boticario (2010), los foros son un servicio de comunicación ampliamente usado en ambientes e-learning por lo que la información obtenida de la interacción en los mismos permite determinar características de colaboración del estudiante, que de acuerdo a Santos et al. (2003), pueden ser representadas como: actividad, iniciativa, y actividad causada por el usuario. Entre los atributos que se consideran indicadores de estas características están: el número de hilos que el estudiante inicia (representa la iniciativa del usuario), el número de mensajes enviados (relacionado con la actividad del usuario), y el número de respuestas a los mensajes enviados por el usuario (relacionado con la actividad causada por el usuario).

Talavera y Gaudioso (2004), deducen de sus resultados que estudiantes con iniciativa son una característica clave para la colaboración y Anaya y Boticario (2010) identifican que los indicadores de la actividad del estudiante y la actividad que es producida por un estudiante en particular en otro estudiante son características claves de colaboración.

En concreto, en un trabajo previo del grupo aDeNu aplicando la metodología propuesta por Santos et al. (2009) basada en técnicas de diseño centrado en el usuario se habían identificado las siguientes recomendaciones relativas al foro:

- Ver contribuciones no leídas del profesor
- Leer contribuciones de un compañero
- Contribuir en el foro para plantear dudas que surjan en el curso

Partiendo de ellas, en esta investigación se ha trabajado en concretar estas recomendaciones y detectar otras que pudieran ser de interés en el uso de la herramienta foros. Para ello, se ha diseñado una experiencia que permita entender

el contexto de uso y recoger datos de la misma para ser analizados. A continuación se ha analizado el contexto de uso y los escenarios en donde se van a ofrecer las recomendaciones. Este proceso ha sido realizado por el profesor, tomando como base las respuestas dadas por sus estudiantes en dicha experiencia por medio de cuestionarios y entrevistas elaborados, además de analizar los datos de interacción. Aplicando técnicas de minería de datos, el profesor ha encontrado oportunidades de recomendación, asociadas con valores concretos de las interacciones de los estudiantes en el curso.

Planteamiento de la experiencia

Nuestra experiencia está centrada en el proceso de extracción de los datos de interacción en la plataforma de aprendizaje OpenACS/dotLRN (Santos et al., 2007), para determinar los patrones de interés en el comportamiento de los estudiantes y así proceder a identificar las situaciones en las que es necesario proporcionar recomendaciones.

Descripción del escenario

Se realizó un experimento con 30 estudiantes del primer ciclo de la carrera de Ingeniería en Sistemas Informáticos y Computación de la UTPL. Del total de estudiante enrolados 29 ingresaron en la plataforma. Esta experimentación fue desarrollada desde el 23 de noviembre hasta el 15 de diciembre de 2009.

Los estudiantes tuvieron acceso al curso de “Fundamentos Informáticos”, en el cuál realizaron las siguientes actividades:

- Interacción con las herramientas disponibles en la plataforma: foros, calendario, noticias, preguntas frecuentes, tareas, documentos compartidos.
- Durante el proceso de interacción, los estudiantes revisaron los archivos y enlaces subidos por el profesor/tutor a la plataforma.
- En base a las lecturas realizadas de los documentos compartidos, desarrollaron dos tareas que luego fueron subidas a la plataforma en la sección “Carpeta de tareas de la comunidad de Fundamentos Informáticos”.
- Participaron de los foros propuestos en el curso.

Obtención de datos de interacción

Finalizada la experimentación, se procedió a la extracción de los datos de interacción de los estudiantes en la plataforma. Para recoger las interacciones de los estudiantes se utilizó la herramienta TAM (Tracking and Audit Module), desarrollada para el framework de OpenACS/dotLRN por el grupo aDeNu. Esta herramienta registra los datos asociados a las acciones del usuario en el sistema, en archivos log que alimentan el modelo de usuario (Couchet et al., 2008). En nuestro caso se recogen las trazas tanto de acciones del usuario en los foros, como de visitas realizadas a los diferentes servicios de la plataforma.

| Atributo | Descripción |
|-------------------------------|---|
| num_sessions | Número de sesiones del usuario |
| average_time_per_session | Tiempo medio por sesión |
| last_visit | Último acceso o visita |
| min_days_from_last_visit | Número mínimo de días desde la última visita |
| max_days_from_last_visit | Número máximo de días desde la última visita |
| num_pages_visited | Número de páginas que el usuario ha visitado |
| num_forums_visited | Número de foros que el usuario ha visitado |
| num_msg_to_group_forum | Número de mensajes que el usuario a enviado a los foros del grupo |
| num_msg_visited | Número de mensajes visitados por el usuario |
| num_average_replies_threads | Número medio de respuestas en los hilos |
| num_threads_user | Número de hilos en los que ha participado el estudiante |
| num_threads_user_in_and_tutor | Número de conversaciones en la que ha participado el usuario y el tutor |
| num_threads_started | Número de conversaciones (hilos) que el usuario ha iniciado. |
| num_folders_visited | número de visitas a las carpetas |
| num_files_visited | Número de archivos visitados |
| num_files_download | Número de archivos descargados |
| num_faq_visited | Número de visitas a las faq |
| num_help_visited | número de visitas a la ayuda |
| num_calendar_visited | Número de visitas al calendario |

Tabla 1. Atributos relacionados a la interacción del usuario con el sistema

Estos datos de interacción del usuario con el sistema (ver Tabla 1) son los que se tomarán en cuenta para la identificación de indicadores para la aplicación de las técnicas de clasificación que permitan obtener los valores de las condiciones

requeridas para el tipo de recomendación de “Leer y Postear mensajes en los foros”, seleccionada en esta investigación.

Además, para complementar los datos de la participación de los estudiantes durante la fase de experimentación con la plataforma OpenACS/dotLRN, se procedió a la aplicación de técnicas de recogida de datos como: encuesta, entrevista y observación, con el objeto de recoger información cualitativa de cuál es el uso de los foros, y que permita contrastar y/o complementar con la que se obtenga del análisis de las interacciones, a fin de poder definir recomendaciones apropiadas al estudiante.

Las variables que se consideraron medir con la aplicación de estas técnicas fueron: 1) Interacción con los foros, 2) Satisfacción de la interacción a través de los foros y 3) Contribución al proceso de aprendizaje del estudiante al interactuar con los foros. La primera variable permitiría validar y complementar los datos recogidos de la interacción con la plataforma. Las otras dos eran necesarias para una próxima fase consiste en determinar si con el uso del sistema recomendador mejora o no el nivel de satisfacción o contribución al proceso de aprendizaje del estudiante.

Resultados de la experiencia

Los resultados obtenidos del procesamiento del archivo de log de la interacción con los foros se pueden observar en la figura 1, donde, de los 29 estudiantes que ingresaron a la plataforma, 28 interactuaron con los foros. El estudiante que resta sólo presentó accesos a la plataforma pero no registró visitas o envíos de mensajes a los foros propuestos.

| Resultados de la interacción en los foros | |
|--|------------------------------|
| Interacción realizada | Número de estudiantes |
| Visitas a los foros | 28 |
| Visitas a los mensajes | 27 |
| Envío de mensajes | 25 |
| Inicio de conversaciones (Hilos) | 17 |
| Envío de mensajes dentro de los hilos | 18 |
| Interacción en conversaciones con el profesor | 11 |
| Respuestas a hilos | 18 |

Figura 1. Resumen de la interacciones con los foros

Como podemos observar, un poco más de la mitad de los estudiantes iniciaron conversaciones en los foros del curso, aunque hay que considerar que no todos los estudiantes se remiten a dar respuesta a las hilos iniciados.

Por otro lado, con las técnicas de recogida de datos cualitativos se determinaron los siguientes aspectos comunes entre los estudiantes:

- Los estudiantes se remiten a dar respuesta a los temas planteados en la plataforma, pocos son los que inician conversaciones nuevas.
- Se observa que los mensajes de algunos de los estudiantes se basan en las respuestas de sus otros compañeros.
- Los estudiantes se expresan libremente a través de los foros, vencen su timidez, tienen una mayor participación, pero algunas de las contribuciones tienen poco fundamento teórico.
- No todos visitan los mensajes ingresados.
- No todos responden a los mensajes comentados por los compañeros.
- No comparten documentos o enlaces de interés, referente a los temas que se proponen.
- En base a la experiencia realizada y al análisis de los datos recogidos con los instrumentos de medición, podríamos decir que es necesario ofrecer recomendaciones cuando:
- El estudiante no ha creado conversaciones.
- No ha ingresado mensajes en las conversaciones iniciadas por los compañeros.
- No ha leído los mensajes ingresado por los compañeros o por el profesor.
- Alguien respondió a su mensaje, lo cual, es conveniente que dé respuesta al mismo.
- No ha compartido documentos o enlaces.

Estas recomendaciones fueron concretadas mediante los valores obtenidos con la aplicación de la técnica de clustering detallada más adelante. De esta forma, se promueve el aprendizaje activo, cuyas ventajas ya se comentaron en la introducción.

Validación de resultados

La validación de los datos recogidos con el TAM se realizó comparando los resultados obtenidos por las tres fuentes: cuestionarios, entrevistas y observación.

A continuación se presenta en la tabla 2 un resumen de la comparación realizada.

| Actividad | Resultados con cuestionario y entrevistas | Resultados procesamiento archivo de log |
|---------------------|---|---|
| Visitas a los foros | 28 | 28 |
| Lectura de mensajes | 24 | 27 |
| Envío de mensajes | 24 | 25 |
| Hilos iniciados | 17 | 17 |

Tabla 2. Cuadro comparativo entre los resultados obtenidos con el cuestionario y mediante el TAM

Con el procesamiento del archivo de log, se obtuvo un número de 27 estudiantes que visitaron mensajes, si este valor lo comparamos con el resultado obtenido con el cuestionario (24 estudiantes que manifiestan realizar lectura de los mensajes antes de ingresar uno nuevo), vemos que existe una mínima diferencia de (3), esto podría deberse a que la formulación de la pregunta en el cuestionario estaba orientada a conocer si el estudiante lee los mensajes de sus compañeros antes de ingresar el suyo. Por otra parte, si consideramos el valor de 24 (resultado con cuestionario) como el número de estudiantes que ingresan mensajes podemos ver que existe una diferencia de una unidad con el valor real obtenido al procesar el archivo de log, ya que este último muestra que 25 estudiantes han enviado mensajes, lo cual no significa una diferencia relevante.

Selección de los patrones de interacción

Para la selección de los patrones de interacción con los que definir las recomendaciones, fue necesario conocer en qué situación o condición es necesario sugerir la recomendación, aspecto que fue complementado con la aplicación de la técnica de clustering.

- *Aplicación de la técnica de clusterig*

Con el objeto de buscar semejanzas en el valor de ciertos atributos que son compartidos por los estudiantes en el momento de la interacción con la plataforma, es necesario trabajar con algoritmos que faciliten esta tarea.

En base a esto y considerando el estudio realizado en el estado del arte se ha creído conveniente trabajar con la técnica de clustering que permite agrupar estudiantes en subclases de acuerdo a su nivel de participación y semejanza de acceso a la plataforma (Mejía et al, 2008).

Con esta técnica se procedió a agrupar a los estudiantes del curso en diferentes grupos relacionados con las actividades realizadas en los foros, y así descubrir patrones que reflejen comportamientos análogos en los estudiantes. Para llevar a cabo este proceso se consideraron los indicadores estadísticos de la interacción de los estudiantes con los foros y mediante la aplicación de la técnica de clustering se pudo observar la relación entre estos indicadores y la participación/colaboración de los estudiantes en esta herramienta.

El conjunto de datos con los indicadores estadísticos de las interacciones de los estudiantes, está formado por:

- Número de sesiones.
- Tiempo medio por sesión.
- Número de visitas a los foros.
- Número de visitas a mensajes.
- Número de mensajes enviados.
- Numero de hilos o conversaciones que el estudiante inicia.
- Número de hilos en los que ha participado el estudiante y el profesor.
- El número de respuestas a los mensajes enviados por el usuario.
- El número de respuestas del usuario a los hilos iniciados.

Estos indicadores hacen referencia a las interacciones activas y pasivas del estudiante en los foros, relacionadas con el estado de colaboración descrito en Santos y Boticario (2004) y con las características de colaboración mencionadas en la sección 3.

Selección del modelo

Se procedió a crear un modelo con la técnica de clustering usando el conjunto de datos tomados de la interacción con los foros. La herramienta seleccionada para esto fue WEKA³ (Witten & Frank, 2005), que es un software de libre distribución de código abierto y que brinda la facilidad de aplicar diferentes técnicas de aprendizaje y minería de datos.

La evaluación del modelado se realizó comparando y analizando los resultados con varias técnicas e interpretando su significado. El proceso fue el siguiente:

Con el algoritmo **SimpleKMeans** se realizaron tres experimentaciones con diferentes valores de semilla (centros iniciales del clúster). En la primera experimentación se consideró el valor por defecto 10, en la segunda se incrementó a 20, y en la tercera a 100, esto con el propósito de ir mejorando los resultados obtenidos. El algoritmo fue configurado para obtener 3 clústeres.

Luego de haber aplicado este algoritmo se realizó una evaluación de los resultados obtenidos en la tres experimentaciones para determinar cuál es el mejor, con lo cual se determinó que los resultados de la primera experimentación no eran muy convenientes para nuestro caso, ya que los grupos eran muy dispares; en la segunda experimentación se obtuvieron mejores resultados, sin embargo la agrupación no era tan clara, ya que los clústeres no quedaban bien definidos; en la tercera experimentación es donde se obtuvieron mejores resultados, con un valor más bajo en la suma de los cuadrados de los errores (11.03).

De igual forma, se realizaron tres aplicaciones del algoritmo **EM** con diferentes valores de semilla, obteniendo resultados similares en las tres experimentaciones, por ello, para la evaluación se presenta la aplicación realizada con los valores por defecto del modelo (semilla 100, clústeres=3), que coinciden con los de la tercera experimentación del algoritmo SimpleKMeans.

| Algoritmo | SimpleKMeans | | | EM | | |
|------------------------------|--------------|----|----|----|----|---|
| | 0 | 1 | 2 | 0 | 1 | 2 |
| Clúster | | | | | | |
| Instancias | 12 | 12 | 6 | 13 | 15 | 2 |
| % de instancias clasificadas | 40 | 40 | 20 | 43 | 50 | 7 |

Tabla 3. Comparación de resultados entre SimpleKmeans y EM

Los resultados que se muestran en la tabla 3 han sido obtenidos considerando los mismos parámetros de configuración para ambos algoritmos: clústeres=3 y semilla=100. Se consideraron 3 clústeres debido al conjunto pequeño de instancias con lo que se contaba y con el propósito de obtener clústeres con valores bajos, medios y altos en los niveles de participación o colaboración de los estudiantes en los foros, lo cual coincide con otras experimentaciones relacionadas (Anaya & Boticario, 2010).

Al analizar en forma individual los valores encontrados para cada atributo con ambos algoritmos, y considerando la distribución de las instancias en cada clúster, se decidió trabajar con los centroides de los atributos obtenidos con SimpleKmeans (ver figura 2), que fue el que presentó una mejor clasificación en su tercera aplicación, ya que los clústeres obtenidos mediante este algoritmo presentaron mayor consistencia y similitud entre sus características. Los valores de estos centroides serán los que se consideren para la definición de las recomendaciones.

```

Clusterer output

kMeans
=====

Number of iterations: 4
Within cluster sum of squared errors: 11.028945683637478
Missing values globally replaced with mean/mode

Cluster centroids:

Attribute                Full Data          Cluster#
                        (30)              (12)              (12)              (6)
-----
num_sessions              5                2.9167           8.5833           2
average_time_per_session 545.5478         343.9653         342.9874         1353.8333
num_forums_visited        13.4333          9.0833           19.25            10.5
num_msg_to_group_forum    1.2333           0                 2.4167           1.3333
num_msg_visited           6.6333           4.3333           9.0833           6.3333
num_average_replies_threads 0.72             0                 1.1333           1.3333
num_threads_started       1.5333           2.0833           0.9167           1.6667
num_threads_user_and_tutor 0.4333           0.3333           0.5833           0.3333
num_answers_to_user_msg   0.8667           0.1667           0.5833           2.8333
num_answered_threads      0.9667           0                 1.8333           1.1667

Clustered Instances

0      12 ( 40%)
1      12 ( 40%)
2       6 ( 20%)
    
```

Figura 2. Tercera aplicación del algoritmo SimpleKMeans – semilla = 100

Así mismo, se presenta un cuadro comparativo de los resultados obtenidos con cada algoritmo para cada una de las características de colaboración de los estudiantes, en los grupos.

| | Clúster 0 | | Clúster 1 | | Clúster 2 | |
|---|----------------|-----------|----------------|-----------|----------------|-----------|
| Algoritmo | <i>k-means</i> | <i>EM</i> | <i>k-means</i> | <i>EM</i> | <i>k-means</i> | <i>EM</i> |
| Nivel de Iniciativa | Alto | Medio | Bajo | Bajo | Medio | Alto |
| Nivel de Actividad | Bajo | Bajo | Alto | Medio | Medio | Alto |
| Nivel de Actividad causada por el usuario | Bajo | Bajo | Medio | Alto | Alto | Medio |

Tabla 4. Comparación de características del estudiante en base a indicadores de colaboración

Como se puede observar en la tabla 4, cada algoritmo obtuvo diferentes niveles en cada caso, no obstante para elegir uno de ellos en particular se tuvo que analizar todo el conjunto de resultados.

Características de los grupos obtenidos con el algoritmo seleccionado

Una vez aplicados los algoritmos, resultaron tres grupos de características distintas, donde considerando los indicadores de colaboración que, de acuerdo a Santos y Boticario (2004), permiten caracterizar a los estudiantes como: a) Participativo: que mide la actividad en diferentes servicios, es decir si el estudiante produce muchas contribuciones (envía muchos mensajes a los hilos de los foros, sus mensajes son respondidos por otros estudiantes, o sólo envían mensajes); b) No colaborativos: contribuye en poco en los foros, es un estudiante que actúa como si no hay colaboración; y c) Con iniciativa: un estudiante que inicia nuevas actividades por sí mismo, envía nuevos mensajes a los foros, inicia nuevas conversaciones, etc. Se procede a describir las características de cada grupo:

Grupo 0: Formado por estudiantes que leen pocos mensajes de los foros (ver figura 2), no realizan envíos de mensajes, es decir realizan muy pocas contribuciones lo cual se ve reflejado en una actividad baja, aunque tienen una alta iniciativa ya que registran la mayor cantidad de hilos iniciados (2,0833); son estudiantes que pasan poco tiempo conectados a la plataforma, sin embargo se conectan más que los del grupo 1 pero menos que los del grupo 2. Podemos decir que son estudiantes que presentan una baja participación en los foros, son poco colaborativos, es decir, se trata de un grupo con una participación baja.

Grupo 1: Estudiantes que envían más mensajes al foro que los otros grupos, leen o visitan alrededor de 9 mensajes, realizan más contribuciones que los del grupo 0 y 2, es decir que tienen una alta actividad, aunque son de baja iniciativa ya que son los que registran el número de hilos iniciados más bajo (0.9167), pasan menos tiempo conectados a la plataforma, aunque el número veces (sesiones) que se conectan a la misma es el más alto (8) que los otros grupos. Así mismo, el nivel de actividad causada por el usuario es medio; por lo que, podemos concluir que se podría decir que se trata de un grupo con una participación general alta.

Grupo 2: Encontramos a estudiantes que se conectan muy poco a la plataforma (número de sesiones=2), menos que los otros grupos, pero con un tiempo medio por sesión más alto, lo que significa que permanecen mucho tiempo conectados en cada sesión. Envían menos mensajes (1.33) que los del grupo 1, por lo que poseen un

nivel medio de actividad y de iniciativa, aunque el nivel de actividad causada por el usuario es alto. Podemos concluir que en este grupo se encuentran estudiantes con mayor participación que los el grupo 0 pero menor que los del grupo 2.

Definición de las recomendaciones

Para la definición de las recomendaciones hay que conocer las condiciones o situaciones en las que es necesario ofrecer las recomendaciones. Estas fueron determinadas teniendo en cuenta el análisis de los datos recogidos con los instrumentos de medición, el procesamiento del archivo de log y el análisis realizado de la aplicación de los algoritmos para los clústeres obtenidos. En el cuadro siguiente se describe las recomendaciones identificadas durante este proceso, con la información obtenida para cada situación mediante las técnicas de clustering.

| Recomendación | Situaciones en la que es necesaria la recomendación | Información obtenida mediante clustering |
|---|---|--|
| <ul style="list-style-type: none"> • Postear un mensaje en el foro compartiendo un documento o un enlace. | <ul style="list-style-type: none"> • Cuando el estudiante no ha compartido ningún documento/ enlace. | El alumno se conecta a la plataforma más de 2 sesiones. |
| <ul style="list-style-type: none"> • Inicie una nueva conversación/hilo. | <ul style="list-style-type: none"> • En caso de que tenga un bajo número de contribuciones. | Número de contribuciones menor a 2. |
| <ul style="list-style-type: none"> • Ingresar mensajes en los hilos iniciados por los compañeros. | <ul style="list-style-type: none"> • Cuando el número de contribuciones del estudiante en el foro es bajo y el hilo tenga una tasa baja de contribuciones (menor a X). | Numero de contribuciones del estudiante menor a 2 y número de contribuciones que hay en el hilo X sea menor a 1. |
| <ul style="list-style-type: none"> • Dar respuesta a los mensajes comentados por los compañeros. | <ul style="list-style-type: none"> • Cuando otro estudiante haya comentado su mensaje. | Número de contribuciones menor a 2. |
| <ul style="list-style-type: none"> • Leer el mensaje del foro propuesto por el profesor. | <ul style="list-style-type: none"> • Cuando aún no haya leído el mensaje propuesto por el profesor en el foro. | Número bajo de contribuciones en el foro, menor a 2. |
| <ul style="list-style-type: none"> • Leer los mensajes de los compañeros. | <ul style="list-style-type: none"> • Cuando no haya visitado los mensajes posteados por los compañeros. | Número bajo de contribuciones en el foro, menor a 2. |
| <ul style="list-style-type: none"> • Visitar las conversaciones/hilos donde se han ingresado más contribuciones. | <ul style="list-style-type: none"> • Cuando un determinado hilo tenga un número alto de contribuciones. | El hilo cuenta con un número alto de contribuciones, igual a 9. |

Tabla 5. Identificación de cuándo es necesario ofrecer una recomendación

Una vez identificadas las situaciones en las que ofrece las recomendaciones, considerando los elementos que se requieren para la definición de recomendaciones mencionados en Santos y Boticario (2008) y en base a lo descrito en el cuadro anterior, se procedió a determinar los atributos del modelo de datos del sistema recomendador que formarían parte de las condiciones de las recomendaciones. Estos atributos con su respectivo valor especifican el valor de las condiciones de aplicación para cada recomendación. Como ejemplo, se muestra la definición de la primera recomendación señalada en el cuadro, siguiendo la plantilla definida para este proceso basada en el modelo de Santos y Boticario (2008).

Antes de proceder al ejemplo, es conveniente mencionar tres aspectos importantes que intervienen en la definición de la recomendación: a) el objeto de la plataforma (en este caso sería el objeto foro), b) la acción (en este caso, leer o postear mensajes) y c) la condición (que en este caso sería la situación o situaciones en la que se debe ofrecer las recomendaciones). La primera fila de la tabla 6 describe la recomendación. En la segunda, se indican en orden el objeto, la acción, el identificador en concreto de dicho objeto en la plataforma y la instancia específica de la plataforma que hace referencia a ese objeto. Las siguientes filas numeradas como Ci hacen referencia a las condiciones que se han definido. En la columna central se describe dicha información según el modelo definido y en la derecha, se explica a alto nivel su significado. Posteriormente se muestra el mensaje que se le ofrecerá al usuario en la interfaz del curso. Debajo se incluye la explicación que acompañará a la recomendación para justificarla. Finalmente, la última línea presenta algunos metadatos que aportan información semántica a la recomendación (en orden): categoría, origen, situación en el curso y relevancia.

| | | | |
|---|--|------------------------------------|---|
| Rec-1: Postear un mensaje en el foro compartiendo un documento o un enlace | | | |
| Recomendar a un estudiante que postee un mensaje en cualquiera de los foros del curso de Fundamentos Informáticos compartiendo un documento o enlace si aún no ha compartido ningún recurso. | | | |
| Forum | contribute | 8057 3323 3324 | UNED_testing |
| C1 | is_in_group= 2287 | | El estudiante debe estar en el curso de Fundamentos Informáticos para que se le dé la recomendación |
| C2 | group_id=2287 object=folder action= contribute pointer = 3102 qualifier=negative | | El alumno no ha subido ningún fichero en la carpeta <u>Documentos compartidos de Fundamentos Informáticos</u> |
| C3 | group_id=2287 object=forum action= contribute | | El alumno ha posteado ya algún mensaje en el foro. |
| C4 | min_num_session_in_plataform =2 | | El alumno ha de haber entrado en la plataforma más de 2 sesiones |
| Puedes postear mensajes en el foro Internet y sus aplicaciones La necesidad pone en peligro la Web Presencia Web 2.0 propuesto en el curso y compartir documentos o enlaces que contribuyan al aprendizaje de tus compañeros. | | | |
| En el foro del curso sólo has ingresado mensajes, sería bueno que compartas también algún recurso con tus compañeros. | | | |
| Soporte técnico | Tutor | Familiarizándose con la plataforma | 5 |

Tabla 6. Ejemplo de definición de una recomendación según el modelo

La validación de las recomendaciones instanciadas ha sido realizada por un educador que ha analizado el modelo de datos propuesto, y está justificada también porque se parte de recomendaciones genéricas que ya venían validadas de una experiencia previa de aDeNu. Lo que interesa ahora validar es la aplicación de dichas recomendaciones en el contexto de uso y ver si suponen una mejora en la experiencia de aprendizaje

Análisis crítico

Luego de aplicado los algoritmos, resultaron tres grupos de características distintas con diferentes niveles para cada una de las características de colaboración de los estudiantes en los foros: actividad, iniciativa, y actividad causada por el usuario (Santos et al., 2003). No obstante, de acuerdo con el estudio realizado por Anaya y Boticario (2010) se puede mejorar la clasificación de los estudiantes a partir de la formación de equipos de trabajo que promuevan el aprendizaje colaborativo

y luego trabajar en base a la relación entre los indicadores propuestos y la colaboración de los estudiantes considerando al foro como espacio de comunicación entre los equipos de trabajo, y así clasificar a los estudiante de acuerdo a su nivel de colaboración usando técnicas de aprendizaje no supervisado que contribuyan a la mejora del proceso de aprendizaje de los estudiantes, de esta forma también se podría administrar mejor los procesos de aprendizaje colaborativos. Estos mismos autores proponen en su trabajo, la construcción de métricas de colaboración usando así mismo técnicas de aprendizaje no supervisado para la asignación de un valor de colaboración a cada estudiante, de tal forma que ellos puedan ser comparados de acuerdo a su colaboración. Si consideramos la posibilidad de trabajar en grupos dentro de los foros para una futura experiencia, las interacciones en los foros podrían ser el soporte para medir la colaboración de los estudiantes.

Se podría también mejorar los grupos obtenidos considerando nuevas características relacionadas con la interacción del estudiante con el sistema, como las mencionadas por Talavera y Gaudioso (2004), para así contar con más conocimiento que permitan contar con un número mayor de grupos y ampliar las recomendaciones a otros servicios de la plataforma.

CONCLUSIONES

El propósito de esta investigación ha sido encontrar valores para las condiciones de aplicación de las recomendaciones, en base a la aplicación de técnicas de aprendizaje no supervisado considerando como fuente de datos las interacciones de los estudiantes en el curso de Fundamentos Informáticos de la carrera de Ingeniería en Sistemas Informáticos y Computación de la UTPL, con el fin de mejorar el soporte adaptativo ofrecido en un escenario de aprendizaje online. Estos estudiantes estuvieron inscritos en un curso en la plataforma OpenACS/dotLRN, cuyo acceso fue proporcionado por los integrantes del grupo aDeNu de la UNED en la cual se encontraba instalada.

Un conjunto de indicadores estadísticos se derivaron de las interacciones de los estudiantes en los foros propuestos, ya que son la principal herramienta para dar soporte al tipo de recomendación de “Leer y Postear mensajes en los foros”, seleccionada para esta investigación.

Una vez determinados los indicadores se tenía que seleccionar la técnica más apropiada para encontrar características similares en el comportamiento de los estudiantes. La técnica seleccionada fue la de clustering. Se realizaron

experimentaciones con el algoritmo SimpleKmeans y EM. Los resultados obtenidos con la ejecución de los algoritmos demostraron que el modelo que mejores resultado presentó fue el obtenido con SimpleKmeans.

En base al análisis de los datos recolectados de la interacción de los estudiantes en la plataforma y a los resultados obtenidos del análisis de la aplicación de los instrumentos de medición, junto con los resultados de la aplicación de los algoritmos, se encontraron que las situaciones en las que es necesario ofrecer las recomendaciones son cuando el estudiante: i) No ha posteado o leído los mensajes, ii) No ha creado conversaciones, iii) No ha compartido documentos o enlaces, iv) No ha visitado los mensajes propuestos por el profesor, o cuando iv) Otro estudiante ha comentado sus mensajes. Identificadas estas situaciones se revisó el modelo de recomendaciones del que se partía y se propuso una serie de recomendaciones a ofrecer en el curso, relativas al uso de los foros. Posteriormente, estas recomendaciones han sido instanciadas formalmente en la plantilla definida para tal efecto, para poderse aplicar en el entorno de aprendizaje.

TRABAJOS FUTUROS

El siguiente paso debe estar enfocado a realizar una segunda experimentación en el mismo contexto pero ahora acompañadas de las recomendaciones identificadas previamente cuando las condiciones tengan lugar. Y así determinar si el uso de recomendaciones mejora el nivel de satisfacción de los estudiantes con los foros y si a la vez mejora su proceso de aprendizaje. Esto se logrará comparando ambas experiencias y/o resultados obtenidos con los instrumentos de medición.

Agradecimientos

Agradecemos a los integrantes del grupo de investigación aDeNu por permitirnos ser parte de las investigaciones hechas dentro de su proyecto y por el apoyo brindado en el desarrollo de esta investigación, agradecemos también a los expertos y usuarios que participaron en las experimentaciones realizadas.

Notas:

- 1 Grupo de investigación aDeNu: <http://adenu.ia.uned.es>
- 2 Proyecto europeo EU4ALL (FP6-2005-IST-5): <http://www.eu4all-project.eu>
- 3 Weka: <http://www.cs.waikato.ac.nz/ml/weka/>

REFERENCIAS BIBLIOGRÁFICAS

- Adomavicius, G.; Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, June.
- Anaya, A.; Boticario, J. G. (2010). Application of machine learning techniques to analyse student interactions and improve the collaboration process. *Expert Systems with Applications: Special Issue on Computer Supported Cooperative Work in Design*. (In Press).
- Boticario, J. G.; Anaya, A. (2009). Clustering Learners according to their Collaboration. In *Proceedings of the 13th international conference on computer supported cooperative work in design (CSCWD 2009)*, IEEE Computer Society Press.
- Couchet, J. ; Santos O.C. ; Raffenne E. ; Boticario J. G. (2008). The Tracking and Auditing Module for the OpenACS Framework, 7th OPENACS / .LRN Conference, Valencia, España. November 18-19.
- Drachsler, H.; Hummel, H.; Koper, R. (2008) Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model. *International Journal of Learning Technology (IJLT)*, Vol. 3, No. 4.
- García, C.; Gómez, I. (2009). Algoritmos de aprendizaje: KNN & KMEANS. [en línea] Disponible en: <http://www.it.uc3m.es/jvillena/irc/practicas/08-09/06.pdf> [consulta 2010, 10 de febrero]
- García, E.; Romero, C.; Ventura, S.; Castro, C. (2009). An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. *User Modeling and User-Adapted Interaction*, Volume 19, 99–132.
- Mejía, C., Mancera, L., Gómez, S., Baldiris, S., Fabregat, R. (2008). Supporting Competence upon DotLRN through Personalization. 7th OpenACS / .LRN conference. Valencia, España. 18-19 November.
- Perera, D.; Kay, J.; Yacef, K.; Koprinska, I. (2007). Mining learners' traces from an online collaboration tool Clustering. *Proceedings of the 13th International Conference of Artificial Intelligence in Education*. Marina del Rey, CA. USA. July.
- Quincey, E.; Kostkova, P.; Farrell, D. (2009). Visualising web server logs for a Web 1.0 audience using Web 2.0 technologies: eliciting attributes for recommendation and profiling systems. In the *Proceedings of the Workshop on Adaptation and Personalization for Web 2.0 in connection with UMAP*, June 22-26.
- Sanjog, R. ; Mahanti, A. (2008). Filler Items Strategies for E@ctive Shilling Attacks. In *Workshop on Recommender Systems*, Patras, Greece. (ECAI).
- Santos, O. C., Boticario, J. G. (2004). Supporting a collaborative task in a web-based learning environment with Artificial Intelligence and User Modelling techniques. 6º Simposio Internacional de Informática Educativa (SIIE'04). November 16-18. Caceres, España.
- Santos, O. C.; Boticario, J. G.; Raffenne, E.; Pastor, R. (2007). Why using dotLRN? UNED use cases. *Proceedings of the FLOSS (Free/Libre/Open Source Systems) International Conference*, 195-212.
- Santos, O.C.; Boticario, J. G. (2008). Users' experience with a recommender system in an open source standard-based learning management system. In *proceedings of the 4yh Symposium on Usability and HCI for Education and Work (USAB)*(in press).

- Santos, O.C., Martin, L., Del Campo, E., Saneiro, M., Mazzone, E., Boticario, J.G., Petrie, H. (2009). User-centered design methods for validating a recommendations model to enrich learning management systems with adaptive navigation support. [en línea] Disponible en: <http://www.easy-hub.org/workshops/umap2009/doc/paper7.pdf> [consulta 2009, 15 de Septiembre].
- Santos, O.C.; Rodríguez, A.; Gaudioso, E.; Boticario, J. G. (2003). Helping the tutor to manage a collaborative task in a web-based learning environment. *Artificial intelligence in education (AIED): Supplementary Proceedings*. Universidad de Sidney, Volume 4, Australia, 153-162.
- Schank, R. C.; Cleary, C. (1995). *Engines for Education*. Lawrence Erlbaum. [en línea] Disponible en: <http://www.engines4ed.org/hyperbook/> [consulta 2009, 15 de Septiembre].
- Sierra, B. (2006). *Aprendizaje Automático: Conceptos básicos y avanzados*. Madrid: Pearson Prentice Hall.
- Talavera, L.; Gaudioso, E. (2004). Mining Student Data to Characterize Similar Behavior Groups In Unstructured Collaboration Spaces. In *Proceedings of the workshop on artificial intelligence in CSCL. 16th European conference on artificial intelligence, ECAI (17-23)*. Valencia, Spain.
- Vialardi, C.; Bravo, J.; Shafti, L.; Ortigosa, A. (2009). Recommendation in Higher Education Using Data Mining Techniques. In *Educational Data Mining (EDM)*.
- Witten, I. H.; Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. 2nd Edition. Morgan Kaufmann.
- Yu, Z.; et al. (2007). Ontology-Based Semantic Recommendation for Context-Aware E-Learning. In *Proceedings of the 4th Conference on Ubiquitous Intelligence and Computing*, v.4611, Berlin, Heidelberg: Springer, 898-907.

PERFIL ACADÉMICO Y PROFESIONAL DE LOS AUTORES

Priscila Valdiviezo Díaz Máster en Educación a Distancia (Universidad Técnica Particular de Loja, Ecuador). Ingeniera en Sistemas Informáticos y Computación (Universidad Técnica Particular de Loja, Ecuador). Docente Investigador, Escuela de Ciencias de la Computación, Modalidad Presencial, Modalidad Abierta y a Distancia, Universidad Técnica Particular de Loja.

E-mail: pmvaldiviezo@utpl.edu.ec

Olga C. Santos Directora Técnica de I+D del Grupo de Investigación aDeNu de la UNED. Diploma de Estudios Avanzados en Inteligencia Artificial (UNED, España). Ha participado -a nivel nacional e internacional- en 12 proyectos de investigación, publicado más de 80 artículos en conferencias y revistas y sido miembro del comité de programa en más de 25 encuentros científicos, 10 de los cuales también como miembro del comité organizador.

E-mail: ocsantos@dia.uned.es

Jesús G. Boticario Dirige el grupo de investigación aDeNu de la UNED aDeNu. Profesor de Informática de la UNED. Coordinador científico en proyectos nacionales y europeos de investigación en el área de e-inclusión y e-learning. Ha participado en 18 proyectos de investigación financiados. Organizador de conferencias y workshops en el área de modelado de usuario en sistemas de aprendizaje. Evaluador de proyectos y artículos de revistas y conferencias nacionales e internacionales. Ha tenido diversos cargos en la UNED en el área de la tecnología aplicada a la educación (p.ej., Vicerrector de Innovación y Desarrollo Tecnológico). Coordinador del proyecto de Accesibilidad y Diversidad Funcional dentro de las Redes de Innovación Docente de la UNED. Miembro del Grupo de eINCLUSION y de eLearning de la Iniciativa Española de Software y Servicios (INES) y de eVIA. Miembro del Consejo Asesor del Centro de Atención a Universitarios con Discapacidad de la UNED (UNIDIS).

E-mail: jgb@dia.uned.es

DIRECCIÓN DE LOS AUTORES

Priscila Valdiviezo Díaz
Universidad Técnica Particular de Loja
San Cayetano Alto.
Loja, Ecuador

Jesús G. Boticario/// Olga C. Santos
Grupo de Investigación aDeNu
ETSI Informática, UNED
Calle Juan del Rosal, 16
Madrid 28040 - España

Fecha de recepción del artículo: 21/04/08

Fecha de aceptación del artículo: 13/10/08