

Herramienta de soporte a la evaluación del aprendizaje y gestión docente

Xavi Canaleta, Xavi Solé, Joan Navarro

Departamento de Informática
La Salle, Universitat Ramon Llull
C/ Quatre Camins 30
08022 Barcelona

{xavic,xsole,jnavarro,dave}@salleurl.edu

Resumen

La presentación de este recurso docente quiere dar a conocer como la docencia universitaria se puede beneficiar de diferentes métodos de Inteligencia Artificial y de técnicas de Minería de Datos para poder cubrir estas necesidades del ámbito educativo. Se han definido funcionalidades afines a las necesidades docentes actuales y se ha desarrollado un software que pretende dar respuesta a cinco casos de uso: predicción del éxito o fracaso del alumno, creación de grupos de trabajo, detección de perfiles de alumnos, descubrimiento de tendencias y visualización de resultados. Este recurso ha sido desarrollado con la implicación del alumnado y su Proyecto Fin de Carrera.

Finalmente, la comunicación presentará el uso que se ha hecho hasta el momento del mismo en el Grado de Ingeniería Informática y de su utilización en otros grados y másteres.

Summary

The presentation of this teaching resource aims to show how university education can benefit from the different methods of Artificial Intelligence and Data Mining techniques to cover the needs of the teaching sector. Functions similar to those of current teaching needs have been defined and we have developed a software targeted at five specific areas: prediction of student success or failure, creation of working groups, detection of student profile, detection of trends and the display of results. This resource has been developed with the collaboration of the students on their Final Year Project.

Finally, this text presents the results of the incorporation of this methodology in the Computer Science Engineering degree programme

and its use in other degree programmes and masters.

Palabras clave

Evaluación del aprendizaje, gestión docente, minería de datos docente, visualización de resultados.

1. Motivación

El Espacio Europeo de Educación Superior (EEES) promueve la instauración de nuevas metodologías docentes en detrimento de las tradicionales clases magistrales. Uno de sus principales objetivos es dar más protagonismo al alumno, haciendo que este sea un actor participe en el proceso de aprendizaje y abandone la figura de receptor pasivo de conocimiento. La evaluación continua y la enseñanza práctica han originado un aumento del uso de sistemas que permiten almacenar información relacionada con el comportamiento de los alumnos y sus resultados. Sin embargo el volumen de información disponible es tan grande que su análisis y estudio tradicionales han devenido impracticables. De este modo ha aparecido la necesidad de aplicar métodos automáticos para la extracción de conocimiento orientado a la comprensión y mejora de los resultados del sistema educativo. Dicho conocimiento puede resultar útil a ambos actores participantes, tanto a responsables académicos como al alumnado.

El desarrollo tecnológico de los últimos años ha dado lugar a la omnipresencia de las Tecnologías de la Información y la Comunicación (TIC) en todos los ámbitos de la sociedad. Los sistemas de información actuales generan y almacenan cantidades descomunales de información.

Actualmente la docencia cuenta con sistemas que permiten registrar datos relacionados con el comportamiento y resultados que obtienen los alumnos. Adicionalmente se dispone de sistemas más innovadores, como sistemas de educación basados en web o sistemas de gestión de contenidos para el aprendizaje. El volumen de datos generado es inmenso y su tratamiento manual resulta impracticable.

El proceso de descubrimiento del conocimiento en bases de datos dota de una solución a este problema. Ofrece marcos de trabajo que permiten transformar los datos de bajo nivel en información útil, es decir, en conocimiento.

La aplicación de estos métodos a las bases de datos del ámbito educativo recibe el nombre de minería de datos docente (*Educational Data Mining*) [7][8]. Actualmente existe una oferta de herramientas para la extracción del conocimiento del sistema docente basado en la web, que centran su objetivo en la obtención de una mayor comprensión del proceso de aprendizaje de los estudiantes, orientado a la mejora de la calidad y el rendimiento del sistema educativo [1][5][6]. Pese a todos estos avances, se observa una falta de herramientas orientadas al sistema educativo tradicional. Existe una oportunidad real para ofrecer sistemas orientados a propósitos específicos de metodologías docentes relacionadas con la adquisición de competencias transversales, la atención personalizada y la implantación de los planes de Bolonia.

Así pues, el primer objetivo del trabajo se centró en realizar una buena definición del sistema a diseñar. Esto implicó un análisis de los modelos de proceso existentes para la extracción de conocimiento y la propuesta de un modelo propio adecuado al entorno de trabajo. Esto condujo a la definición de los casos funcionales de aplicación de los métodos de descubrimiento de conocimiento en el ámbito docente y la elaboración de la especificación de requisitos software de la herramienta a desarrollar.

El objetivo final del proyecto era la producción de un software la ambición del cual es convertirse en una aplicación de código libre que permita, a usuarios no expertos en el campo de la minería de datos, obtener información útil, válida y original de los datos procedentes del sistema docente.

2. Funcionalidades del sistema

Antes de realizar la descripción de las funcionalidades de la herramienta, merece la pena destacar que se realizó un exhaustivo trabajo para obtener una especificación de requisitos software de alta calidad. Esto fue posible gracias al esfuerzo de un Trabajo Final de Máster [9] que, posteriormente, permitió a través de un Proyecto Final de Carrera [10] desarrollar el sistema.

Para llegar a la obtención de los requisitos software de alta calidad primero se estudió el proceso de descubrimiento de conocimiento en bases de datos [3]. Seguidamente se realizó el análisis de los modelos de proceso [3][4] y se definió un modelo de proceso propio. Una vez se hubieron estudiado los problemas y las técnicas más representativas del aprendizaje automático [2] se pasó a definir los casos funcionales de aplicación de los métodos de descubrimiento de conocimiento. Finalmente se realizó la especificación de requisitos del sistema.

El sistema, a partir de uno o varios ficheros de calificaciones de los alumnos, permitirá llevar a cabo diferentes funcionalidades. El formato de los ficheros de entrada del sistema puede obtenerse fácilmente desde cualquier programa que permita el procesamiento de hojas de cálculo, como por ejemplo Microsoft Excel o Calc de OpenOffice. El formato de los ficheros es CSV (*Comma Separated Values*) delimitado por el separador punto y coma, con la particularidad de diferenciar la primera línea del resto. La primera línea comprende el nombre de las columnas o de los atributos que almacena el fichero. Las dos primeras etiquetas no pueden cambiarse y son orientadas a la identificación del alumno, estas son: Expediente y Nombre. A continuación figura el nombre de tantas etiquetas como calificaciones o notas se disponga, ya sean estas de ejercicios, prácticas, exámenes, etcétera. Estas son de carácter alfanumérico, pueden contener espacios y deben aparecer ordenadas según tienen lugar en el tiempo, de menos reciente a más. Las siguientes líneas del fichero comprenden los datos de los alumnos. Cada valor, separado por punto y coma, hace referencia al atributo correspondiente especificado en la primera línea del fichero. Los valores de la columna Expediente pueden comprender valores alfanuméricos y deben ser únicos, es decir, no pueden figurar alumnos distintos con el mismo identificador. En el campo

Nombre deben aparecer los apellidos y el nombre de los alumnos separados por coma. Las demás columnas, dedicadas a las calificaciones de las distintas pruebas, pueden comprender valores numéricos decimales (con el separador coma) o bien las constantes NP (no presentado) y MH (matrícula de honor).

A partir de estos ficheros se pueden realizar las siguientes funcionalidades:

- Predicción del éxito o fracaso de los alumnos. Con esta opción el profesor facilitará al sistema dos ficheros de datos. El primero será un histórico de notas de uno varios cursos anteriores. El segundo serán las calificaciones disponibles del curso actual hasta la fecha. El sistema dará como resultado una predicción de las calificaciones que obtendrán los alumnos según la experiencia que ha suministrado el profesor mediante el histórico de notas.
- Creación de grupos de trabajo. Mediante esta funcionalidad el profesor suministrará al sistema datos con las calificaciones de los alumnos que se desea agrupar. Seguidamente configurará las características de los grupos que quieren obtener y el sistema dará como resultado una propuesta de configuración de grupos que satisfaga los criterios introducidos. Adicionalmente, para facilitar la comprensión de la naturaleza de los grupos de trabajo generados, el sistema complementará la propuesta con elementos visuales como gráficos radiales y gráficos de barras.
- Detectar perfiles de alumnos. Con esta funcionalidad el profesor facilitará al sistema información con notas de los estudiantes de los cuales se quiere detectar perfiles, configurará el número mínimo y máximo de perfiles a detectar, podrá fijar la relevancia de los datos suministrados y el sistema dará como resultado un conjunto de perfiles de alumnos con características similares. El sistema, para facilitar la lectura de las características de los perfiles, mostrará diferentes elementos gráficos, como gráficos circulares o histogramas.
- Descubrimiento de tendencias. Esta opción permitirá al profesor obtener de los datos que ha facilitado al sistema las posibles relaciones existentes entre las diferentes calificaciones obtenidas por el alumno durante el curso académico. Estas tendencias permitirán al

docente poder analizar la relevancia y adecuación de las actividades diseñadas dentro de la asignatura.

- Análisis de los resultados de los alumnos. Mediante esta funcionalidad el profesor podrá analizar y comparar, al mismo tiempo, los resultados que han obtenido los alumnos en dos actividades de evaluación. El sistema mostrará un conjunto de valores estadísticos, histogramas, gráficos de pastel, gráficos de correlaciones y descripciones textuales para su interpretación.

3. Plataforma tecnológica

A continuación se hace una breve descripción de las diferentes tecnologías utilizadas en el desarrollo del proyecto:

- Weka (*Waikato Environment for Knowledge Analysis*) es una plataforma para la aplicación de métodos de minería de datos, escrita en Java, bajo la licencia GPL (*General Public License*). Esta plataforma se ha usado tanto a nivel de usuario de la aplicación, para analizar y comprender las técnicas aplicadas, como a nivel de programador, modificando ciertas clases para adaptar los algoritmos necesarios en el desarrollo del sistema.
- Se ha decidido utilizar el paradigma de programación orientada a objetos y concretamente el lenguaje de programación Java.
- Por lo que se refiere a la programación de las interfaces gráficas, se ha optado por la librería SWING de Java. Esta forma parte de las *Java Foundation Classes*.
- Para cubrir las necesidades de exportación de los resultados a Excel, se ha decidido utilizar la librería Java Excel API Java.
- Como herramienta de modelado se ha escogido para poder analizar, diseñar y documentar el sistema el lenguaje de modelado unificado UML (*Unified Modeling Language*) usando la aplicación StarUML.

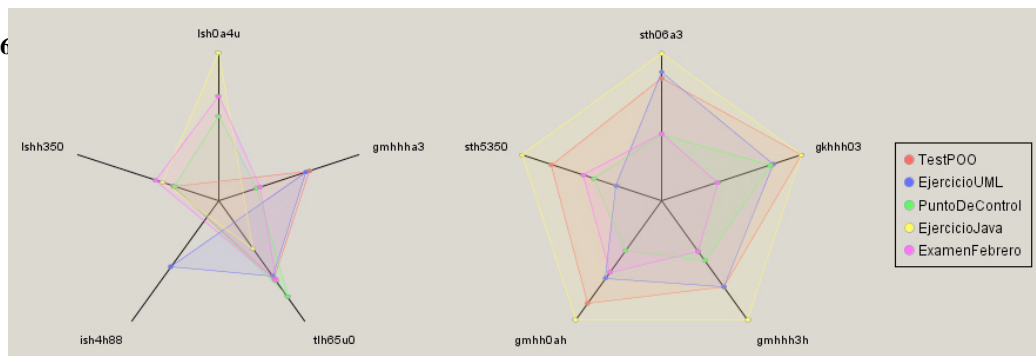


Figura 1. Gráficos radiales de ejemplo de visualización de grupos heterogéneos (izquierda) y homogéneos.

4. Aportaciones y utilidad

Una vez descritas las diferentes funcionalidades del sistema, en este apartado se pretende presentar la utilidad que conllevan los diferentes casos de uso implementados en la aplicación.

Por lo que se refiere a la predicción del éxito o fracaso del alumno, esta funcionalidad da soporte al profesor para poder estimar la evolución de su alumnado a mitad de un semestre. Esta herramienta, además de disminuir su carga de gestión docente y permitirle focalizar su esfuerzo en la evaluación del aprendizaje, es una aliada excelente de cara a acciones de tutoría que se deban realizar con el alumno.

La creación de grupos de trabajo es una tarea cada vez más habitual en las metodologías de aprendizaje cooperativo y en el aprendizaje basado en proyectos. Las aportaciones de esta funcionalidad son diversas:

- La automatización de dicha tarea supone una disminución clara del tiempo de dedicación del profesor a este proceso.
- La herramienta permite la creación tanto de grupos homogéneos como heterogéneos de cualquier número de personas. Esto hace que sea muy flexible y aplicable a cualquier materia de los estudios de Ingeniería Informática.
- El soporte gráfico adicional que proporciona dicha funcionalidad, facilita el análisis y visualización de los grupos creados. Como se muestra en la Figura 1, el profesor tiene una visión clara del grado de homogeneidad o heterogeneidad de cada uno de los grupos

propuestos por el sistema gracias a los gráficos radiales, por ejemplo.

La detección de perfiles de alumnos tiene muchas posibilidades de utilización, aunque algunas de las que se mencionan aún no hayan sido totalmente explotadas:

- La detección de perfiles permite un conocimiento más profundo del grupo clase con el que se trabajará. Esto permite un mejor diseño de estrategias de aprendizaje viendo la tipología diversa (o no) del alumnado.
- El sistema está preparado para poder realizar esta operación basándose tanto en calificaciones de las diferentes actividades de evaluación realizadas (las clásicas notas) como en la gradación de las diferentes competencias evaluadas. Así pues, la generación de perfiles puede permitir obtener información sobre el perfil competencial del alumnado de la clase.
- También cabe destacar aquí el soporte gráfico para hacer más intuitivo el análisis de los resultados obtenidos, tal y como se puede apreciar en la Figura 2.

El descubrimiento de tendencias es una funcionalidad que genera diferentes reglas de asociación entre diversas actividades que han sido evaluadas y calificadas a los alumnos. La manipulación necesaria, para obtener reglas de asociación o tendencias, realizada de forma manual sería totalmente inasumible a nivel de coste temporal. Pero, además, los resultados pueden permitir al profesor detectar ciertos errores en el diseño de actividades de evaluación y realizar mejoras. Por ejemplo, analizando las

reglas obtenidas podremos ver qué notas condicionan o están directamente relacionadas con otras obtenidas posteriormente. Pero, a la vez, podemos detectar que las calificaciones en ciertas actividades no tienen ninguna implicación con otras realizadas posteriormente. Así debería cuestionarse si realmente esa actividad cumple la función por la cual ha sido diseñada. Y, de este modo, llegar a la conclusión que ésta deba ser cambiada para que tenga relevancia en el proceso de aprendizaje.

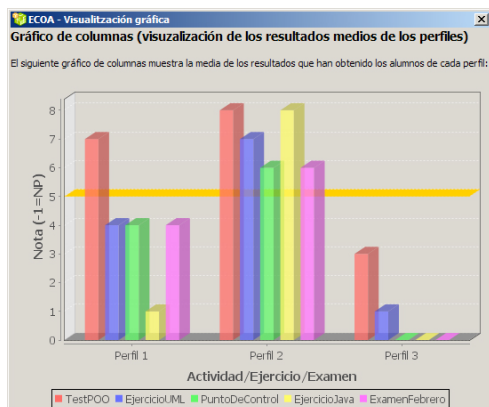


Figura 2. Gráfico de columnas con los resultados de los tres perfiles detectados.

Finalmente, la funcionalidad que permite el análisis de resultados de los alumnos (la única que no implica el uso de técnicas de aprendizaje automático para su implementación) ha sido una de las funcionalidades de mayor uso en el poco tiempo que el sistema lleva en producción en diversos ámbitos de la escuela universitaria. Las aportaciones más relevantes que ofrece son:

- La visualización gráfica de los resultados de cualquier actividad de evaluación que se facilita al sistema, mediante elementos como los gráficos de pastel, los histogramas, etc.
- El análisis de la correlación entre los resultados obtenidos entre dos actividades y su interpretación visual de dicho análisis, tal y como muestra el ejemplo de la Figura 3.

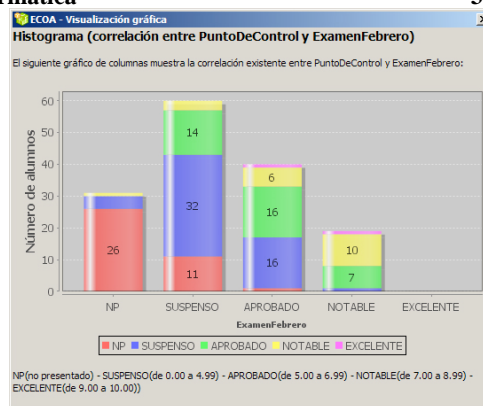


Figura 3. Ejemplo de histograma de correlación de los resultados de dos actividades.

5. Resultados y conclusiones

El impacto del software aquí presentado dentro de la Escuela de Ingeniería se puede considerar moderado hasta el momento dado que este estuvo disponible en una versión productiva a finales de septiembre de 2011. En este corto periodo de tiempo ha sido usado en diferentes estudios y ámbitos. A continuación se detallan los más significativos:

- Asignatura de Álgebra de primer curso del Grado de Informática. Basa gran parte de su metodología docente en el aprendizaje basado en problemas y en el trabajo cooperativo. Ha usado la herramienta para la generación de grupos de trabajo heterogéneos.
- Asignatura de Programación de primer curso del Grado de Informática: ha utilizado la aplicación para visualizar los resultados del examen del *midterm*, así como para analizar la correlación entre las calificaciones de evaluación continua y los resultados obtenidos en dicha prueba.
- Asignatura de Programación y Estructura de Datos de segundo curso del Grado de Informática: uso del sistema para la generación de los grupos de prácticas y también para verificar los perfiles de los alumnos del grupo-clase.
- Asignatura de Sistemas Operativos de tercer curso del Grado en Informática: análisis de resultados obtenidos por el alumnado y su

correlación entre las calificaciones obtenidas en prácticas y las obtenidas en los exámenes de problemas y conocimientos.

- Finalmente, Máster Universitario de Formación del Profesorado: uso experimental del sistema para analizar los diferentes perfiles propuestos basados en las calificaciones competenciales de los alumnos durante el primer semestre en los módulos genérico y específico, tanto para alumnos de la especialidad de Tecnología, Idiomas o Educación Física. También se usó en algunas materias del Máster para generar los grupos de trabajo heterogéneos para ciertas actividades.

El *feedback* recibido por parte del profesorado de estas asignaturas ha sido completamente positivo, aportando sugerencias en cuanto a flexibilizar aún más ciertos aspectos de adquisición de los datos por parte del sistema.

Como líneas inmediatas de trabajo está la presentación de la herramienta al resto del profesorado en un taller interno de innovación docente, así como la publicitación de la aplicación a otras facultades de universidad. De hecho, la presentación de este recurso docente en las Jornadas de la Enseñanza Universitaria de la Informática (JENUI 2012) responde a este afán de expansión del sistema.

El aplicativo ECOA es propiedad de La Salle, Universitat Ramon Llull. La intención es distribuirlo a toda la comunidad universitaria, que esta pueda hacer uso del mismo y, de este modo, poder recibir *feedback* y seguir evolucionando el sistema. El software se encuentra en las dos páginas web personales de los autores (<http://users.salleurl.edu/~xavic> en el apartado de material y <http://users.salleurl.edu/~xsole>) donde se puede descargar libremente.

Referencias

- [1] Avouris, N., et al. *Why logging of fingertip actions is not enough for analysis of learning activities*. s.l.: In Workshop on usage analysis learning systems at the 12th international

Recursos docentes

- conference on artificial intelligence in education, 2005.
- [2] E. Shapire, R. *COS 511: Foundations of Machine Learning*. s.l.: Princeton University Computer Science Department, 2003. Lecture 1.
- [3] Fayyad, U., Piatetsky-Shapiro, G. y Smyth, P. *From Data Mining to Knowledge Discovery in Databases*. Providence: American Association for Artificial Intelligence, 1996. p. 37-54. Vol. FALL 1996. 0738-4602-1996.
- [4] J. Cios, K. et al. *Data Mining - A Knowledge Discovery Approach*. New York : Springer, 2007. 978-0-387-33333-5.
- [5] Mazza, R. y Milani, C. *Exploring usage analysis in learning systems: Gaining insights from visualisations*. s.l.: In workshop on usage analysis in learning systems at 12th international conference on artificial intelligence in education, 2005.
- [6] Merceron, A. y Yacef, K. 2005. *Tada-ed for educational data mining*. s.l.: Interactive Multimedia Electronic Journal of Computer-Enhanced Learning, 2005.
- [7] Romero, C y Ventura, S. *Educational data mining: A survey from 1995 to 2005. Expert Systems with Applications*. s.l.: Elsevier, 2006, Vol. Expert Systems with Applications.
- [8] Romero, C., Ventura, S. y García, E. *Data mining in course management systems: Moodle case study and tutorial*. s.l.: Elsevier science, 2007.
- [9] Solé-Beteta, X. *Definició del sistema d'extracció de coneixement de l'àmbit docent ECOA*. La Salle, Universitat Ramon Llull, Barcelona, 2011. Trabajo Final de Máster. Directores: Orriols-Puig, A. y Canaleta, X.
- [10] Solé-Beteta, X. *Desenvolupament del sistema d'extracció de coneixement de l'àmbit docent ECOA*. La Salle, Universitat Ramon Llull, Barcelona, 2011. Proyecto Final de Carrera. Directores: Orriols-Puig, A. y Canaleta, X.