

# En busca del estudiante perdido: caracterización de los ‘no presentados’

Ismael Sanz, María J. Aramburu, Lledó Museros, María Pérez, Carmen Barrachina

Departamento de Ingeniería y Ciencia de los Computadores

Universitat Jaume I

Avg. Vicent Sos Baynat s/n

12071 Castellón de la Plana

{isanz,aramburu,museros,mcatalan,carmen.barrachina}@uji.es

## Resumen

El abandono de asignaturas por parte de los estudiantes es un problema cada vez más importante. Este trabajo describe una metodología y una herramienta informática para entender qué lleva a los estudiantes a abandonar una asignatura, y así poder realizar acciones preventivas y paliativas.

Nuestro objetivo inicial era ser capaces de determinar empíricamente unos criterios de clasificación que permitieran dar pistas a priori sobre el riesgo de abandono de cada estudiante matriculado, y así poder poner en marcha medidas preventivas desde el principio del curso. Se han realizado pruebas piloto en asignaturas del ámbito de las Bases de Datos, con resultados prometedores.

El trabajo realizado nos condujo al desarrollo inicial de un sistema de información integrado, con la flexibilidad suficiente para servir de apoyo a futuras tareas de tomas de decisiones relativas a la gestión de la docencia.

## Palabras clave

Abandono estudiantil, sistema de apoyo a la decisión.

## 1. Introducción

### 1.1. Contexto y motivación

La motivación inicial del presente trabajo surgió de la experiencia acumulada durante la realización de tres proyectos de mejora educativa, dos de ellos asociados a la asignatura Bases de datos avanzadas (optativa de cuarto curso de Ingeniería Informática y de tercer curso de Ingeniería Técnica en Informática de Gestión), y el otro asociado a la asignatura

Bases de datos (obligatoria de segundo curso de Ingeniería Técnica en Informática de Gestión en modalidad semipresencial). Estos proyectos consistieron globalmente en la implementación de una nueva metodología de enseñanza centrada en el estudiante y basada en un sistema de evaluación continua. Todos los proyectos alcanzaron su objetivo principal, que era mejorar el número de aprobados, de manera incluso más satisfactoria de lo esperado. Sin embargo no cambió en absoluto, e incluso empeoró, otra de las variables que nos habíamos propuesto mejorar: el número de estudiantes presentados.

Desde nuestro punto de vista, habíamos conseguido que quien llevara la asignatura al día superara la asignatura, pero al precio de una tasa alta de abandonos. Esto se podría achacar a la nueva metodología que se implantó. Sin embargo, entendemos que la situación no es tan simple: cada curso encontramos estudiantes matriculados que no llegan a tener ningún contacto con la asignatura (no consta un solo acceso al curso correspondiente en el Aula Virtual, no asisten ni siquiera a la clase inicial de presentación, ni se presentan a los exámenes). Naturalmente, entre estos dos extremos (estudiantes que por cualquier circunstancia abandonan la asignatura tras intentarlo, y estudiantes que simplemente no aparecen) existe toda una serie de casuísticas intermedias.

Como un último intento de paliar esta situación, en la asignatura Bases de Datos Avanzadas incluimos un sistema de evaluación adicional. Consistía en un itinerario específico para estudiantes que no pudieran realizar los entregables realizados en clase, bajo la suposición implícita de que la imposibilidad de asistir a las clases presenciales era el motivo principal de las cifras de abandono; hay que tener en cuenta que por las características de estas asignaturas muchos de los estudiantes matriculados están realizando simultáneamente su estancia en prácticas

o trabajando en una empresa. Los estudiantes que eligieran este itinerario debían realizar unas tareas previamente programadas, disponían de tutorías físicas (en el despacho del profesor) y virtuales (por e-mail o a través del Aula Virtual), y tenían la obligación de presentarse al examen final de la asignatura. Sin embargo, la experiencia ha mostrado que este itinerario no es seguido por prácticamente ningún estudiante, y que muchos estudiantes que encuentran trabajo o están haciendo la estancia en prácticas no tienen problema alguno en seguir la evaluación continua.

Debía haber, pues, otras causas subyacentes que desconocíamos. Conversaciones informales con los estudiantes y con profesores en cuyas asignaturas ocurre el mismo problema nos convencieron de que no podíamos plantear nuevas propuestas de soluciones paliativas sin antes descubrir esas causas subyacentes. Nuestra sensación era que muchos estudiantes abandonaban nuestras asignaturas y no sabíamos por qué, ni podíamos hacer nada para evitarlo, lo que no nos parecía aceptable. En una búsqueda bibliográfica en algunos de los principales foros de investigación en innovación educativa encontramos pocos trabajos que arrojaran luz sobre los “no presentados”, que se interpretan más bien como un hecho inevitable – algo con lo que no podemos estar de acuerdo sin una completa investigación previa que no parece haberse realizado hasta ahora [2][6].

Decidimos que una buena manera de abordar esta investigación era recopilar datos suficientes para permitir estudiar el problema utilizando herramientas informáticas de análisis multidimensional de datos. Se trata de herramientas de inteligencia de negocio de amplia utilización para el apoyo de decisiones en entornos empresariales y financieros, hemos considerado que sería muy interesante utilizarlas para analizar nuestra problemática y de esta manera realizar una investigación mucho más completa que la que permiten otras herramientas más simples pero de uso mucho más habitual, como pueden ser las hojas de cálculo. Además, disponer de una herramienta con estas características podría llegar a ser muy útil para múltiples tareas relacionadas con la docencia.

Una aproximación similar había sido ya ensayada por [7], que mostró la utilidad de aplicar análisis estadísticos sobre información almacenada en un *Data Warehouse* específicamente construido para estudiar el absentismo en la Universidad Industrial de San-

tander. Este trabajo se basaba exclusivamente en la información disponible en el sistema de información universitario, e ignoraba otros factores referentes a la situación personal y las percepciones subjetivas de los estudiantes que, sospechábamos, resultaban decisivos en el abandono o no de una asignatura. Como veremos, esta hipótesis de trabajo se ha visto confirmada por los resultados obtenidos.

Para asegurarnos de que nuestros resultados fueran lo más generalizables posible, decidimos realizar un trabajo coordinado entre tres asignaturas que comparten área de conocimiento (las bases de datos), pero que se imparten de manera muy diferente: *Bases de datos* (IG18S) es una asignatura semipresencial y armonizada con el EEES; *Bases de datos avanzadas* (IG52/II52) se evalúa, como se ha comentado, de forma continua, mientras que *Sistemas de información inteligentes* (II73) está basada en clases magistrales y se evalúa en un examen final. Por tanto, de esta manera podemos tener en cuenta la manera de impartir la asignatura como un factor más de nuestro análisis.

## 1.2. Objetivos

El objetivo general de nuestro trabajo es obtener información sobre los estudiantes que abandonan las asignaturas, para ser capaces de entender qué les lleva a ello y poder realizar acciones preventivas y paliativas en el futuro. Según nuestra experiencia y la de otros trabajos relacionados [2][6], entendemos que es muy improbable que todos los casos respondan a las mismas causas, y que es muy importante que seamos capaces de descubrir estas causas para entender mejor a nuestros estudiantes y así favorecer su progresión. Por esta razón, nuestro método consiste en la implementación de herramientas para intentar *agrupar* los estudiantes que comparten ciertas características comunes y así *caracterizar* cada uno de esos distintos “grupos de riesgo”, que probablemente necesitarán soluciones específicas. De esta manera sería posible determinar unos criterios de clasificación que permitan determinar *a priori* el riesgo de abandono de un estudiante matriculado, y poder así poner en marcha medidas preventivas desde el principio del curso.

Por tanto, los objetivos de este trabajo pueden desglosarse de la siguiente manera:

- La recopilación de un conjunto de datos que

incluyan las características potencialmente relevantes de los estudiantes matriculados en las asignaturas objeto del proyecto (con datos sobre su situación laboral, académica y personal), debidamente mantenidos en soporte electrónico en forma de base de datos, y pre-procesados para respetar el anonimato de los estudiantes (anonimización).

- La implementación y aplicación de un conjunto de herramientas de análisis de esos datos que nos permita (i) caracterizar a los estudiantes que han abandonado la asignatura, encontrando características que los separen de los que han continuado hasta el final (ii) encontrar los grupos de estudiantes no presentados que presentan características similares (es decir, los “grupos de riesgo”) y (iii) estudiar si hay diferencias entre las distintas maneras de impartir asignaturas.

Un objetivo adicional, que fue creciendo en importancia conforme avanzábamos en el trabajo, es el diseño y la implementación de un sistema de información integrado de apoyo al docente, y la evaluación de su potencial utilidad. Este sistema integraría toda la información académica de las asignaturas y de los estudiantes matriculados, y proporcionaría herramientas adecuadas para su análisis multidimensional. Como resultado obtendríamos una herramienta de apoyo a la toma de decisiones para docentes y gestores académicos.

### 1.3. Resultados

Como resultado de este trabajo hemos creado una infraestructura básica consistente en:

1. Un estudio sistemático de las variables que pueden influir en el abandono de una asignatura.
2. Métodos que permiten obtener datos relativos a esas variables; hemos implementado métodos para extraer información de las bases de datos corporativas de la universidad, y también hemos creado una encuesta electrónica para los estudiantes.
3. Técnicas para producir modelos predictivos, que permiten determinar el grado de riesgo de cada estudiante.
4. Una herramienta de análisis multidimensional de datos que permite estudiar los datos des-

de diversas perspectivas y realizar agrupaciones específicas a las necesidades concretas de cada caso de estudio.

Los puntos 2, 3 y 4 constituyen el embrión de un sistema de información de apoyo para el docente.

El resto del artículo se organiza de la siguiente manera: la sección 2 presenta la metodología desarrollada para la consecución de los objetivos anteriormente planteados, en la sección 3 se realiza un análisis de los resultados obtenidos y finalmente en la sección 4 se exponen conclusiones, se repasa el impacto que están teniendo los resultados preliminares obtenidos y se plantean ideas para trabajos futuros.

## 2. Metodología desarrollada

Nuestra hipótesis inicial de trabajo es que hay dos tipos de factores fundamentales que afectan en la decisión de que un estudiante abandone una asignatura: por un lado, factores relativos a la situación académica de cada estudiante (por ejemplo, el número de créditos superados o el número de convocatorias a que se ha presentado ya) y, por otro, otros relativos a las circunstancias ajenas (por ejemplo, laborales y familiares) y a preferencias personales del estudiante respecto a la asignatura (por ejemplo, el interés que le genere la metodología docente). Deberemos, por tanto, recopilar información sobre ambos aspectos.

Es importante tener en cuenta desde el principio el diferente carácter de ambos tipos de factores: en el primer caso se trata de datos objetivos, de los que podremos disponer de forma exacta y numérica. Sin embargo, los factores personales de los estudiantes son mucho más genéricos y abiertos: preferencias, circunstancias personales, etc. Esto motivará que necesitemos utilizar estrategias de análisis diferenciadas para cada uno de estos factores.

En todo caso, a partir de la hipótesis, será necesario decidir qué información es potencialmente de interés para el análisis, y cómo es posible acceder a ella. Dado el carácter abierto de esta investigación, intentamos determinar un amplio abanico de variables potencialmente interesantes. Como paso inicial se realizaron tres acciones:

- Una búsqueda bibliográfica en foros de innovación educativa. El trabajo más cercano que encontramos fue [6], en el que se describe una

encuesta realizada en la universidad de Oviedo en un proyecto de innovación similar a éste.

- Entrevistas informales con estudiantes (antiguos y actuales) de las asignaturas objeto del proyecto.
- Un estudio de la información disponible sobre el historial académico de cada estudiante en el sistema de información interno de la universidad.

A partir de estos datos, diseñamos la siguiente estrategia:

- Para analizar la información relativa al historial académico de los estudiantes, se aplicaron técnicas de clasificación automática, para intentar determinar la relevancia de cada variable y comprobar si era posible encontrar patrones relevantes para cada asignatura.
- Para analizar las circunstancias personales y las preferencias de cada estudiante, mucho más complejas, optaremos por un análisis exploratorio de los datos, mediante un interfaz gráfico apropiado.

Como algoritmos de clasificación automática se seleccionaron dos métodos sencillos de generación de árboles de decisión: C4.5 [4] y la función genérica *rpart* disponible en el software estadístico R [5], que sigue las técnicas descritas en [1]. Dado que el volumen de datos disponibles no era elevado, no cabía esperar un gran poder predictivo de los modelos obtenidos; sin embargo, este primer paso proporcionaría información muy precisa sobre la relevancia de los factores académicos en cada asignatura.

Para estudiar los factores relativos a las circunstancias y las preferencias personales, diseñamos una encuesta anónima a partir de [6], y usamos el software QlikView [3] para visualizar, explorar y analizar los datos de manera interactiva. La Figura 1 muestra una de las pantallas desarrolladas para el análisis.

### 2.1. Obtención de los datos

Utilizamos la información recopilada para adaptar el contenido de la encuesta propuesta por [6] y generar nuestra propia encuesta. En resumen, se tuvieron en cuenta factores académicos, indicadores del rendimiento, circunstancias personales, y otros factores inherentes a la asignatura (horarios, relevancia para el estudiante...) y al profesorado.

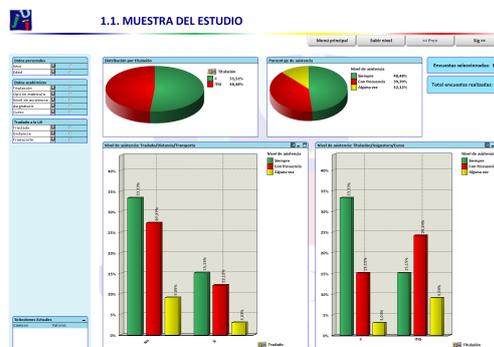


Figura 1: Análisis de resultados en QlikView

Además de diseñar la encuesta, se buscaron otras fuentes de datos de las que podíamos obtener información relevante, a partir de la información disponible en el sistema de información interno de la universidad. En particular, esta información incluye los resultados de cursos anteriores (incluyendo los estudiantes repetidores) y los comentarios de los estudiantes realizados en años anteriores como parte de la evaluación del profesorado.

Una vez determinada qué información debíamos capturar directamente de los estudiantes, en esta fase completamos el diseño de las encuestas a realizar y la automatización de la obtención de datos. Para la implementación de la encuesta consideramos prioritario facilitar lo más posible a los estudiantes su cumplimentación, lo que implicaba los siguientes principios: (1) la encuesta debía ser completamente anónima, para evitar cualquier sospecha y (2) la encuesta debía implementarse mediante una plataforma electrónica tan fácil de usar como fuera posible. El primer principio descartaba el uso del Aula Virtual de la universidad, ya que requiere autenticación. Por ello, realizamos un estudio de varias plataformas de encuestas electrónicas gratuitas disponibles en Internet. Finalmente nos decidimos por *e-encuesta*, ya que proporcionaba la mayor simplicidad de uso a los estudiantes. Se decidió enviar a los estudiantes el enlace tras la finalización de las revisiones de examen de la convocatoria de junio. La figura 2 muestra una pantalla de dicha encuesta.

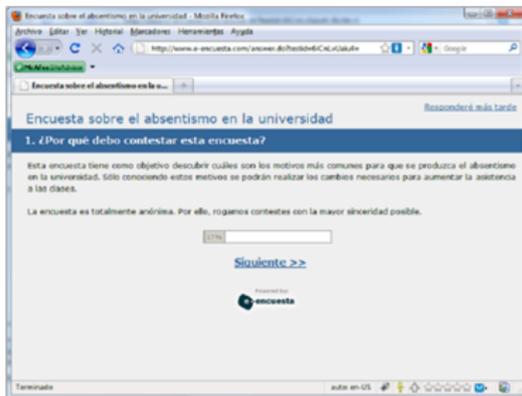


Figura 2: Pantalla de bienvenida de la encuesta

### 3. Puesta en marcha y resultados

En cuanto a la información obtenida del sistema de información interno de la universidad, para su extracción se realizaron los siguientes pasos para cada asignatura involucrada y para cada curso disponible:

1. Extraer la lista de estudiantes del curso.
2. Eliminar los apartados con información personal y los no relevantes. Los campos considerados relevantes son: *Curso*, *Asignatura*, *Convocatoria*, *Titulación*, *Créditos superados*, *Créditos troncales superados* y *Media del expediente*.
3. Añadir información sobre si el estudiante se ha presentado o no (codificada como presentado=1 y no presentado=0). Esto se obtuvo de: (i) para el curso actual, la información de seguimiento de la asignatura, y (ii) para los cursos anteriores, de las actas.

Los datos obtenidos fueron los siguientes: 80 registros asociados a la asignatura II73 (desde el curso 2005/2006), 184 registros asociados a la asignatura IG52/II52 (desde el curso 2007/2008, en el que cambió el sistema de evaluación de la asignatura) y 59 registros asociados a la asignatura IG18S (desde el curso 2008/2009, curso en el que comenzó a impartirse de manera semipresencial). En total se han utilizado 323 registros.

Una vez obtenida toda esta información se comenzó con su análisis.

#### 3.1. Análisis de datos

Por un lado se analizaron los datos obtenidos sobre la situación académica de los estudiantes, y por otro lado, y de manera paralela, los datos obtenidos con la realización de la encuesta, para lo cual se utilizó la herramienta de análisis multidimensional de datos QlikView.

El objetivo del análisis de los datos provenientes del sistema de información interno de la universidad fue identificar, de entre las variables disponibles, cuáles son relevantes para determinar si un estudiante se ha presentado finalmente o no. Para ello, aplicamos las dos técnicas antes mencionadas, que intentan determinar la relación entre una variable aleatoria y diversos factores. En este caso, la variable objetivo era si el estudiante se presentó o no a la asignatura; los factores son el resto de variables disponibles. A partir de este análisis se obtuvieron los siguientes resultados:

- Modelo predictivos para II73. El particionamiento recursivo obtuvo un modelo muy simple (figura 3a). Si la media del expediente era inferior a 6.485 clasifica al alumno como No presentado, y si la media es igual o superior como presentado. La clasificación es correcta en un 90%. Es notable que la única variable relevante para este caso encontrada por el algoritmo ha sido la media del expediente. Intuitivamente parece un resultado razonable, ya que al tratarse de estudiantes que están prácticamente acabando la carrera hay poca variación en el resto de variables.
- Modelos predictivos para IG52/II52 (figura 3b). Como en el caso anterior, depende básicamente de la media del expediente (siendo aquellos estudiantes con media más alta más probable que se presenten, y aquellos con la media más baja menos probable). Sin embargo, en el rango intermedio de puntuaciones entran en juego también el número de créditos troncales superados. El error de clasificación es alto (25%), lo que apunta al hecho de que hay más factores determinantes además de los estudiados.
- Modelos predictivos para IG18S (figura 3c). La partición recursiva obtiene un modelo en el que la variable principal es la media del expediente y secundariamente se utilizan los datos de créditos superados.

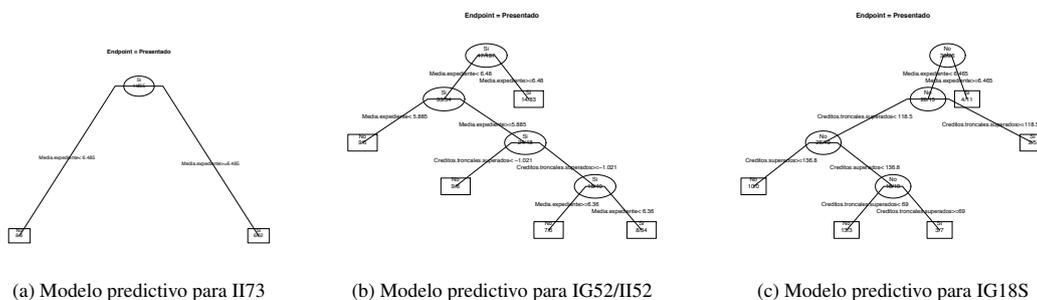


Figura 3: Modelos predictivos desarrollados

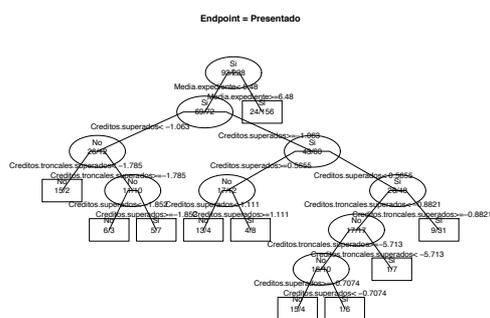


Figura 4: Modelo predictivo para los datos agregados

- Modelos predictivos para los datos agregados (figura 4). A pesar de presentar un error de predicción superior al 30%, son interesantes para mostrar tendencias globales (los datos se han normalizado para limitar las diferencias entre las asignaturas).

A continuación se detallan las conclusiones globales derivadas del análisis de los resultados de las encuestas a través de la herramienta QlikView, resumidas en la figura 5.

En primer lugar, los motivos de absentismo en relación a los horarios de las asignaturas son:

- *Solapamiento de horarios con otras asignaturas*: Una posible solución consistiría en comprobar entre qué asignaturas se producen mayores solapamientos e intentar evitarlos.

- *Muchas clases seguidas*: El hecho de tener muchas clases seguidas sin descanso entre ellas provoca muchas veces que el estudiante decida dejar de asistir a alguna de ellas para poder descansar. Se debería intentar dejar espacios de tiempo entre clases de asignaturas, al menos de las que pertenecen al mismo curso.

- *Muy temprano*: Cabría pensar que esto lo opinen los estudiantes que residen fuera de la ciudad y que deben desplazarse hasta la universidad, requiriendo esto un tiempo determinado. Sin embargo, los que han contestado que las clases son demasiado temprano, en su mayoría residen en la misma ciudad en la que se encuentra la Universidad.

En segundo lugar, los motivos de absentismo más comunes obtenidos de las preguntas abiertas nos dicen que:

- *Organización*: Estos temas son los más problemáticos a la hora de poder asistir a clase, sobre todo lo referente al solapamiento de asignaturas. Se podrían tomar las mismas medidas mencionadas anteriormente.
- *Motivos de salud*: El siguiente motivo más común es el relacionado con motivos de salud, pero en este caso no hay nada que la parte del profesorado puede hacer para solucionar este problema.
- *Solapamiento con el trabajo*: Al igual que antes, a pesar de tratarse de otro motivo común, el profesorado no puede intervenir.
- *Clases aburridas*: En este caso el profesorado sí que puede tomar medidas: intentando impartir clases que sean del interés del estudiante, en las

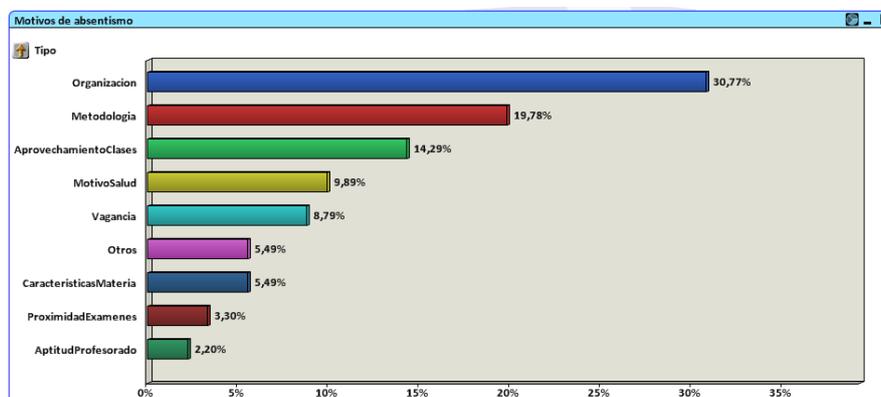


Figura 5: Resumen de los motivos de absentismo manifestados en la encuesta

que éste sea más partícipe, clases más dinámicas en las que se evite entrar en una monotonía que pueda interferir en la atención del alumno.

- *Horario adecuado*: El motivo más popular entre los encuestados hace referencia a tener un horario adecuado de las clases, que no comience demasiado temprano ni acabe demasiado tarde, que no sea en horarios de comida.
- *Puntuar asistencia*: El hecho de premiar la asistencia es importante para los estudiantes, ya que con ello tienen un incentivo extra que les “obligue” a acudir a las clases.
- *Buenas explicaciones*: Otro hecho importante es que el profesorado ofrezcan buenas explicaciones durante sus clases.
- *Ejercicios prácticos*: El mismo nivel de importancia le dan los alumnos a las buenas explicaciones y a la realización de ejercicios prácticos durante las clases.

En cuanto a los motivos más comunes para abandonar una asignatura son los siguientes:

- *Difícil*: El hecho de que la asignatura sea demasiado difícil es el factor más importante a la hora de abandonar la asistencia a una materia. Sería conveniente intentar enfocar las clases de manera que la comprensión de los conceptos de esa asignatura, complicada de por sí, sea más sencilla.
- *Descuelgue*: A veces se produce un descuelgue de la asignatura, debido a la falta de tiempo para llevarla al día, o por la dificultad de ésta.

- *Falta de tiempo, exceso de asignaturas*: Estos dos motivos tienen el mismo nivel de importancia para los encuestados. Es, incluso, posible que estén relacionadas entre ellas, ya que un exceso en la matriculación de asignaturas provocará que el alumno no disponga de suficiente tiempo para atenderlas todas.

Finalmente, en relación a la proporción de presentados en las asignaturas bajo estudio en este trabajo, las conclusiones obtenidas son las siguientes:

- Respecto a *créditos superados*: Los alumnos que más se presentan al examen de alguna de estas asignaturas son los que tienen entre el 51 y el 75 % de créditos superados.
- Respecto a *créditos troncales superados*: Los alumnos que más se presentan al examen son los que tienen entre el 51 % y el 75 % de créditos troncales superados.
- Respecto a *la nota media de expediente*: Los alumnos con notable como nota media del expediente tienen un índice mayor de presentados.
- Respecto a *la titulación y a la asignatura*: Los alumnos se presentan más a la asignatura IG52/II52 que al resto.

Estas últimas conclusiones confirman las conclusiones obtenidas con el análisis de los datos del sistema de información interno a través de la herramienta R.

#### 4. Impacto, conclusiones y trabajo futuro

El primer resultado de este proyecto ha sido la creación de una infraestructura para la gestión integrada de la información referida a la docencia que no está limitada a la aplicación concreta que acabamos de describir. A medio plazo, esperamos convertirla en un completo “Sistema de Análisis de Información Académica” que pueda servir de apoyo para la toma de decisiones tanto a los docentes responsables de las asignaturas como a los coordinadores académicos que gestionan el grado. A tal efecto, nuestra línea de trabajo principal va a estar orientada a la integración de la información disponible en el sistema de información institucional de la universidad, con el objeto de explorar las posibilidades que así se abrirán.

En cuanto al objetivo que motivó inicialmente este trabajo, la caracterización de los no presentados, las técnicas creadas nos han permitido aprender más sobre las causas que llevan a algunos estudiantes a abandonar las asignaturas, refinar los modelos predictivos y el sistema de visualización y análisis mediante QlikView, y aplicar su resultado a la mejora docente. Es importante notar que nuestros resultados indican claramente que cualquier sistema de información útil para el apoyo al docente no puede limitarse a contener sólo información extraída automáticamente del sistema de información institucional, sino que se debe enriquecer en la medida de lo posible con datos estadísticos que sean indicadores de la situación personal y las percepciones subjetivas de los estudiantes, y que dependen de cada materia.

Esto implica necesariamente sistematizar la realización de encuestas a los estudiantes. En la primera aproximación realizada, quisimos imponer el anonimato total en las encuestas, lo que hizo imposible cruzar los datos recogidos con los obtenidos del sistema de información interno de la universidad, limitando enormemente las posibilidades de análisis automático. Por ello, en este curso hemos querido mejorar el sistema de encuestas para poder incrementar los datos disponibles para su análisis automático, limitando las encuestas a los temas más “sensibles” en los que el anonimato tenga sentido.

A pesar de estas limitaciones, los datos obtenidos durante el curso pasado ya han motivado algunos cambios relevantes en la docencia. En la asignatura

Sistemas de información inteligentes (la única que ha acabado en el momento de enviar esta contribución) se han aplicado algunos pequeños cambios en la metodología docente motivados por los resultados de las encuestas, que han redundado en una mejora clara: este curso ningún estudiante ha abandonado la asignatura.

#### Agradecimientos

Este trabajo se ha realizado gracias al apoyo de la *Unitat de Suport Educatiu* de la Universitat Jaume I (a través de proyectos de mejora educativa en los cursos 2009/10 y 2010/11) y del Departamento de Ingeniería y Ciencia de los Computadores de la Universitat Jaume I, que ha financiado ayudas para Carmen Barrachina e Ismael Sanz a través de su Plan Estratégico.

#### Referencias

- [1] Breiman, L.; Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Wadsworth & Brooks/Cole Advanced Books & Software
- [2] Agustín Cernuda del Río, Sonia Hevia Vázquez, María del Carmen Suárez Torrente, Daniel Gayo Avello, “Un estudio sobre el absentismo y el abandono en asignaturas de programación”. XIII Jornadas de Enseñanza Universitaria de la Informática (2007).
- [3] <http://www.qliktech.com>
- [4] Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.
- [5] <http://www.r-project.org/>
- [6] Rodríguez, R. et al. “El absentismo en la Universidad: resultados de una encuesta sobre motivos que señalan los estudiantes para no asistir a clase”. Aula Abierta, 82, ISSN 0210-277, págs. 117-146 (2003)
- [7] Salazar, A.; Gosálbez, J.; Bosch, I.; Miralles, R.; Vergara, L.; A case study of knowledge discovery on academic achievement, student desertion and student retention. 2nd International Conference on Information Technology: Research and Education, UTRE 2004 (2004).