

UNA PROPUESTA METODOLÓGICA DE ESTUDIO DE RENDIMIENTO ACADÉMICO EN UNA TITULACIÓN UNIVERSITARIA

Jorge Más-Estellés, José V. Benlloch-Dualde, M^a José Vicent-López, M^a Ángeles Calduch-Losa, Rosa Alcover-Arándiga, José M. Valiente-González

E.T.S. Informática Aplicada

Universidad Politécnica de Valencia

Resumen

En este trabajo se presenta una propuesta metodológica para analizar los datos de rendimiento académico de los estudios de Informática en ocho universidades españolas, para las tres titulaciones actualmente vigentes (Ingeniería Informática, Ingeniería Técnica en Informática de Gestión e Ingeniería Técnica en Informática de Sistemas). Se estudian, de forma diferenciada, los alumnos de nuevo ingreso y el total de alumnos matriculados en la titulación, analizándose indicadores como alumnos matriculados, ingresados, egresados, demandas en primera opción, abandonos, tasa de éxito, tasa de rendimiento y duración de los estudios. Para todos estos indicadores se calculan unos valores de referencia, promediando los datos disponibles de todas las universidades participantes.

Para cada Universidad, se estudian, durante cinco cursos académicos, los flujos de alumnos en cada titulación. A partir de estos diagramas, considerando solamente aquellas titulaciones con un régimen estabilizado de alumnos (sin cambios de plan de estudios), se representa un diagrama de flujo agregado para cada titulación informática.

En todos los casos, los alumnos están desagregados por grupos, lo que permite observar la influencia que tienen, en el rendimiento académico, las distintas características analizadas.

Con vistas a posteriores estudios utilizando técnicas de minería de datos, se valora el proceso de obtención de datos.

Palabras Clave: *Rendimiento académico, indicadores, Informática, perfil alumnos, estudio descriptivo.*

1. Introducción

El rendimiento académico es un asunto al que, desde hace algún tiempo, se le está dedicando una atención creciente. Los trabajos que tratan de cuantificar el nivel académico del alumnado o de los Centros se llevan a cabo, bien de forma periódica y estructurada (Informes PISA, pruebas e informes institucionales del MEC o de los Gobiernos autonómicos, PNECU, ...), o bien de forma aislada y esporádica, fundamentalmente a iniciativa de grupos de profesores que dan a conocer sus estudios en ponencias o artículos en Congresos o revistas especializadas.

Esta creciente preocupación por el rendimiento académico viene determinada por factores tanto de índole económica como política y social. Cada vez es más discutido el que un estudiante de una Universidad pública permanezca año tras año en la Universidad ocupando una plaza que tiene un elevado coste para la sociedad en términos económicos. También en términos de estricta justicia social, pues esa plaza podría ser ocupada por otro alumno que no pudo acceder a esos estudios y que, quizá, podría haber tenido un aprovechamiento mayor. También en términos sociales, la permanencia de los estudiantes en la Universidad durante prolongados períodos de

tiempo es un aspecto controvertido y causante, en parte, de la cada vez mayor demanda de las titulaciones de primer ciclo frente a las de segundo ciclo (1). Por otra parte, los futuros procesos de acreditación de los nuevos títulos universitarios de grado y master, así como los propios modelos de financiación de las Universidades, tendrán muy en cuenta que la duración efectiva de los estudios se ajuste en buena medida a la contemplada en sus planes de estudios, lo que explica el creciente interés que suscita el tema del rendimiento académico.

El rendimiento académico puede ser abordado desde distintas perspectivas, tanto en lo que a aprovechamiento por parte del alumnado se refiere, como a resultados obtenidos por los distintos Centros, con el fin de diagnosticar tanto a unos como a otros y tomar las decisiones que correspondan. En cualquier caso, está comúnmente aceptado que el conocimiento de la situación en términos numéricos es el primer paso para la propuesta de medidas que puedan corregir situaciones indeseadas.

Los parámetros utilizados para estos estudios pueden ser muy variados, pero cada vez más se están consolidando algunos que, además de resultar relativamente sencillos de calcular, dan una idea bastante fiel de la situación que se intenta estudiar. En este sentido, el Ministerio de Ciencia e Innovación (M.I.C.I.N.N) recoge (2) una lista de indicadores de rendimiento del sistema universitario.

La utilidad de estos estudios, en una primera instancia, es meramente descriptiva, con el fin de obtener unos valores de rendimiento que permitan conocer la situación. Estos valores que se obtienen no tienen, per se, una significación absoluta, por lo que es útil establecer comparaciones con Centros de similares características, no con ánimo competitivo sino cooperativo, para establecer unos valores de rendimiento que puedan considerarse normales, buenos, inaceptables, etc... En una segunda instancia, estos valores de rendimiento calculado pueden utilizarse para identificar los factores que, en mayor medida, influyen en ellos: características socioeconómicas, edad, estudios previos del alumnado y de su entorno al inicio de sus estudios, actividad, o no, laboral de los estudiantes durante sus estudios, características organizativas y docentes de los Centros, planes de estudios, métodos evaluativos, etc.... Todos ellos son factores que podrían, en gran medida, explicar el éxito o fracaso de un estudiante, de modo que la Universidad pudiera incidir sobre ellos. En realidad, algunas Escuelas y Universidades ya organizan actividades en este sentido, como los programas tutoriales (programa PATU de la UPV), o cursos cero de prerequisites. No obstante, el carácter generalmente indiscriminado de estas actividades creemos que resta efectividad a los mismos. Probablemente, si tuviéramos una herramienta que permitiera identificar "a priori" carencias y necesidades de los alumnos en función de sus características, podríamos incidir en ellas con mayor efectividad.

Una metodología ampliamente utilizada en estudios sobre rendimiento es la de hacer encuestas a los distintos estamentos del sistema educativo. Esta metodología tiene el inconveniente de reflejar opiniones subjetivas que además suelen estar muy influenciadas por la amplia mitología existente sobre cuáles son las causas del fracaso escolar, incluso aunque no contemos con parámetros objetivos que apoyen estas ideas. En contrapartida, las encuestas tienen en cuenta parámetros difícilmente cuantificables en cuanto a metodología y características organizativas que de otro modo son de difícil consideración.

Una alternativa a la realización de encuestas es la investigación directa en las bases de datos que recogen información tanto sobre las características e historial del alumnado, como de la Universidad. Para ello es imprescindible contar con una base de datos que recoja una amplia información durante un período de tiempo suficientemente largo, algo que depende, fundamentalmente, de la política que cada Universidad haya llevado a cabo en este sentido, y que esta información se pueda manejar, siempre respetando la legislación vigente en cuanto a protección de datos. A partir de toda esta información, los autores de este trabajo hemos construido almacenes de datos independientes de la base de datos transaccional, con el objetivo de poder trabajar cómodamente. Indagando en este almacén de datos hemos hecho algunos estudios al respecto utilizando técnicas multivariantes, de modelización estadística y de minería de datos (3-9). De este modo, buscamos relaciones entre factores que figuran en la base de datos, poniendo de relieve dependencias y obteniendo modelos que nos permitan predecir el rendimiento o las dificultades que tendrán determinados alumnos según las características propias de cada uno.

Estas técnicas se pueden aplicar utilizando distintas aplicaciones informáticas, tanto comerciales como por ejemplo Clementine (10), como de libre disposición, como por ejemplo Weka (11).

La utilización de Minería de Datos requiere el conocimiento detallado de las características del alumnado, algo que no siempre es posible. Los datos existentes en cuanto a números de créditos aprobados y suspensos, y tasas de éxito y rendimiento suelen presentarse de forma agregada, por lo que es difícil correlacionar estos datos con algunas de las características de los alumnos. El conocimiento detallado de datos desagregados (alumno por alumno) resultaría de gran utilidad para estudios más precisos que relacionaran el rendimiento con distintas características socio-económico-educativas del alumnado. No obstante, y pensando en un estudio que abarcara distintas Universidades, con el fin de obtener conclusiones que no se circunscriban a un ámbito geográfico reducido, pensamos que pueden existir problemas para conseguir este tipo de datos de forma generalizada, debido fundamentalmente a la preservación de la confidencialidad de los datos o a políticas de tratamiento de la información propias de cada Universidad, aunque también pensamos que esas dificultades pueden ser soslayadas.

La metodología que aquí se propone, se circunscribe a estudios descriptivos que abarcan varias universidades, sin entrar a utilizar estas técnicas de Minería de Datos. Aprovechando todo el proceso de captación de datos, sí que se puede valorar la posibilidad de acometer estudios de Minería de Datos para varias universidades en una segunda etapa.

2. Objetivos

El objetivo de este trabajo es presentar una propuesta metodológica para estudiar el rendimiento académico de una determinada titulación universitaria en un conjunto de Centros, con la idea tanto de establecer unos valores de referencia que nos permitan saber si el rendimiento académico de un determinado centro se puede considerar alto, bajo, o similar al de otros Centros, como de conocer qué características de los alumnos influyen en su rendimiento.

El trabajo que aquí se propone se ha aplicado a los estudios de Informática (titulaciones de Ingeniería Informática, Ingeniería Técnica en Informática de Gestión e Ingeniería Técnica en Informática de Sistemas) en ocho Centros españoles. La financiación del trabajo ha estado a cargo del Ministerio de Ciencia e Innovación (antiguo M.E.C.), en el marco del programa de Estudios y Análisis (EA2007-0152, convocatoria 2007). Los resultados de este trabajo no han sido aún publicados por el Ministerio, por lo que aquí presentamos tan sólo su planteamiento y metodología, que pensamos puede ser extensible a otras titulaciones y universidades.

3. Información solicitada e indicadores calculados

Pensando que la obtención de determinada información podría plantear dificultades técnicas, hemos reducido nuestro estudio a la obtención de unos pocos indicadores, intentando que esta información tenga un mínimo nivel de desagregación, como puede ser la diferenciación de los alumnos según su sexo, edad y cupo por el que accede a los estudios, y nota de ingreso. Toda esta información debe, en principio, ser suministrada por cada Universidad mediante sus bases de datos o información de los propios Centros. También se tiene en cuenta la demanda de cada titulación en primera opción, así como el número de estudiantes que finalmente ingresa. Cada universidad aporta la información que considera oportuna o que es capaz de reunir; en algunos casos corresponde al conjunto de varios centros que imparten las titulaciones informáticas, y en otros, sólo a algún centro de la universidad.

En cuanto a los indicadores a calcular, planteamos un estudio a dos niveles:

3.1 Alumnos de nuevo ingreso en la titulación

En un primer nivel, estudiaremos indicadores referidos al primer año de estudios de los alumnos (alumnos de nuevo ingreso), ya que somos conscientes de la especial problemática e importancia

que el primer año de estudios plantea. Aspectos relacionados con el historial de los nuevos alumnos tienen, durante este curso, una incidencia que no resulta tan evidente en cursos posteriores. Los indicadores a estudiar se refieren a las tasas de éxito, rendimiento, tasa de abandono tras el primer año de estudios, y también cálculo del rendimiento académico teniendo en cuenta no sólo el número de créditos superados, sino también las calificaciones obtenidas, y la convocatoria en la que se superó la asignatura. Tanto la información como los indicadores calculados a partir de ella, se refieren a tres cursos distintos: los cursos 2003-2004, 2004-2005 y 2005-2006.

Con estas premisas, cada una de las universidades participantes, para cada una de las titulaciones informáticas ofertadas, y para cada uno de los tres años estudiados, debía completar una tabla como la que se muestra a continuación:

TABLA DATOS ALUMNOS DE NUEVO INGRESO						
UNIVERSIDAD						
TITLACION	II <input type="checkbox"/>		ITIG <input type="checkbox"/>		ITIS <input type="checkbox"/>	
CURSO ACADEMICO	2003-04 <input type="checkbox"/>		2004-05 <input type="checkbox"/>		2005-06 <input type="checkbox"/>	
CREDITOS MATRICULADOS POR ALUMNO						
	MATRICULADOS	DEMANDA 1ª	TASA EXITO	TASA RDTO.	TASA ABAND.	RDTO *
GLOBAL						
HOMBRES						
MUJERES						
EDAD <19						
19<=EDAD<=25						
EDAD>25						
COU/LOGSE/PAU						
FP						
>25 AÑOS						
TITULADOS						
EXTRANJEROS						
Nota Ing >=5 y <6						
Nota Ing >=6 y <7						
Nota Ing >=7 y <8						
Nota Ing >=8 y <9						
Nota Ing >=9						

Tabla 1 Información recabada sobre los alumnos de nuevo ingreso en una titulación

En esta tabla, en la columna DEMANDA 1ª figurarán todos los alumnos del perfil correspondiente demandantes de la titulación en 1ª opción.

TASA DE ÉXITO es la relación entre el número de créditos aprobados y el de presentados en alguna convocatoria.

TASA DE RENDIMIENTO es la relación entre el número de créditos aprobados y el de matriculados.

TASA DE ABANDONO es la relación entre el número de alumnos que abandonan los estudios tras el primer año, y los matriculados ese mismo año.

RDTO (Rendimiento) es un indicador que trata de “afinar” algo más la tasa de rendimiento. Tiene en cuenta no sólo el número de créditos aprobados, sino también la calificación obtenida y la convocatoria en que se aprobó. Se calcula mediante una fórmula que tiene en cuenta los aspectos mencionados, y ya ha sido utilizado en estudios previos (5,7).

3.2 Total de alumnos de la titulación

En un segundo nivel, estudiaremos indicadores relacionados con el conjunto total de alumnos de la titulación. Estos indicadores estarán referidos a cada titulación de cada universidad en cinco cursos diferentes: desde 2001-2002 hasta 2005-2006.

Toda esta información se resume en una tabla como la siguiente:

TABLA TOTAL DE ALUMNOS MATRICULADOS EN LA TITULACION					
UNIVERSIDAD					
TITULACION	II <input type="checkbox"/>		ITIG <input type="checkbox"/>		ITIS <input type="checkbox"/>
CURSO	2001-02 <input type="checkbox"/>	2002-03 <input type="checkbox"/>	2003-04 <input type="checkbox"/>	2004-05 <input type="checkbox"/>	2005-06 <input type="checkbox"/>
	MATRICULADOS	INGRESADOS	EGRESADOS	ABANDONOS	DURACION
GLOBAL					
HOMBRES					
MUJERES					
EDAD Ing < 19					
19 <= EDAD <= 25					
EDAD Ingr > 25					
COU/LOGSE/PAU					
FP					
> 25 AÑOS					
TITULADOS					
EXTRANJEROS					
Nota Ing >= 5 y < 6					
Nota Ing >= 6 y < 7					
Nota Ing >= 7 y < 8					
Nota Ing >= 8 y < 9					
Nota Ing >= 9					

Tabla 2 Información recabada sobre el total de alumnos de una titulación

En esta tabla, todas las columnas se dan en número de alumnos, salvo la columna DURACION, que se refiere a la duración media de los estudios (en años), y es dada por cada universidad de acuerdo con sus cálculos.

Es necesario aclarar que se han considerado ABANDONOS los casos de los alumnos que, sin haber concluido los estudios, no aparecen matriculados al curso siguiente.

EGRESADOS se ha considerado a los alumnos que han completado el total de créditos de la titulación, independientemente de que administrativamente hayan solicitado, o no, el título.

4. Metodología de trabajo y representación de resultados

El proyecto se coordina desde una de las universidades (en este caso, la Universidad Politécnica de Valencia). Cada una de las universidades recopila, ordena y elabora los datos solicitados, organizándolos en las tablas citadas, y los remite a la universidad coordinadora, junto con aquellas características metodológicas y organizativas propias de cada Centro. La universidad coordinadora, manteniendo el contacto con el resto de universidades, filtra y homogeneiza los datos recibidos, aclarando posibles errores, y elabora los resultados globales y referidos a cada grupo de población; tras ello se fija una reunión conjunta de todos los participantes con el fin de obtener las conclusiones del trabajo. Estas conclusiones quedan plasmadas en el informe final.

En la universidad coordinadora, toda la información recibida de las universidades fue volcada en una base de datos implementada en la plataforma Oracle. El motivo fue poder utilizar la herramienta Discoverer para diseñar un almacén de datos con una estructura adecuada al

contexto del estudio que pretendemos realizar y permitir así extraer la información necesaria de manera sencilla, mediante informes dinámicos.

Los indicadores obtenidos son relacionados tanto con las características disponibles de los alumnos, como con las características organizativas y metodológicas propias de cada Centro. Estos indicadores sirven, también, para establecer unos valores de referencia en cada una de las tres titulaciones; no se trata de establecer un "ranking" de Universidades, sino de poner información a disposición de los Centros para que puedan proponer las medidas que cada uno de ellos considere oportunas.

Ante la ingente cantidad de información disponible, hemos optado por hacer dos tipos de representaciones diferentes, según tratemos con los alumnos de nuevo ingreso, o con el total de alumnos de la titulación. Eso no significa que se puedan hacer las mismas representaciones para unos que para otros, pero por la distinta problemática que presentan unos y otros, y también por brevedad, hemos optado por ello.

En cuanto a los alumnos de nuevo ingreso, pensamos que en su rendimiento tienen primordial importancia sus características sociales y su historia previa al ingreso en la universidad. Por ese motivo, proponemos una representación de gráficos de barras de cada uno de los indicadores de estos alumnos en la que aparezcan valores relativos de cada grupo de alumnos. En los gráficos de total de alumnos matriculados, matriculados en 1ª opción, hombre y mujeres, estos valores relativos representan porcentajes sobre el total de matriculados. En el resto de indicadores, se han sumado los valores de cada indicador para todos los grupos de alumnos, y ese valor se ha tomado como 100.

Este tipo de representación tiene el inconveniente de que, para conocer el valor absoluto del indicador de un grupo determinado de alumnos, es preciso consultar la tabla de datos, pero por el contrario, da una visión comparativa muy clara y ajustada de ese indicador para los distintos grupos de alumnos.

Como ejemplo, en la figura 1 se muestra la representación de la tasa de éxito, para cada grupo de los alumnos de nuevo ingreso en la titulación de Ingeniería Informática en una de las universidades participantes, en el curso 2003-2004. En la tabla 3 se pueden ver los valores correspondientes a la gráfica.



Figura 1. Representación de la tasa de éxito de alumnos de nuevo ingreso en II en el curso 2003-2004 por grupos de alumnos

Como ejemplo, en esta representación se puede ver como la tasa de éxito de hombres y mujeres es muy parecida (0,80 frente a 0,78), por lo que en estas gráficas, cada grupo contribuye a la tasa de éxito total con el 50 % aproximadamente.

También como ejemplo, por edades, los mayores de 25 años (amarillo) tienen una tasa de éxito de 1, probablemente debido a que sólo hay dos individuos. No obstante, frente al resto de tasas de éxito de los otros grupos de edades (0,81 y 0,63), los mayores de 25 años representan aproximadamente el 40 % de la tasa de éxito total.

	MATRICULADOS	TASA ÉXITO
GLOBAL	148	0,80
HOMBRES	122	0,80
MUJERES	26	0,78
EDAD <19	130	0,81
19<=EDAD<=25	16	0,63
EDAD>25	2	1,00
COU/LOGSE/PAU	138	0,80
FP	0	
>25 AÑOS	2	1,00
TITULADOS	0	
EXTRANJEROS	5	0,56
Sin Nota Ing	0	
5<=Not Ing<=6	6	0,86
6<=Not Ing<=7	41	0,67
7<=Not Ing<=8	59	0,80
8<=Not Ing<=9	34	0,85
Nota Ing >9	8	0,95

Tabla 3. Valores correspondientes a la figura 1

Esta representación se puede hacer para cada indicador, y también calcular los valores medios ponderados (con el número absoluto de alumnos) de cada indicador en el conjunto de las universidades estudiadas, para cada grupo de alumnos, y hacer representaciones similares de dichos valores medios, que nos dan una idea del rendimiento general de una titulación.

En cuanto al total de alumnos matriculados en una titulación, proponemos una representación de diagramas de flujo de una titulación, donde se muestra, para los años estudiados y para el total de alumnos matriculados en ella, el flujo de alumnos, con las entradas y salidas de alumnos correspondientes. Como ejemplo, en la figura 2 mostramos el gráfico de flujo correspondiente a la titulación de Ingeniero Informático en una de las universidades participantes.

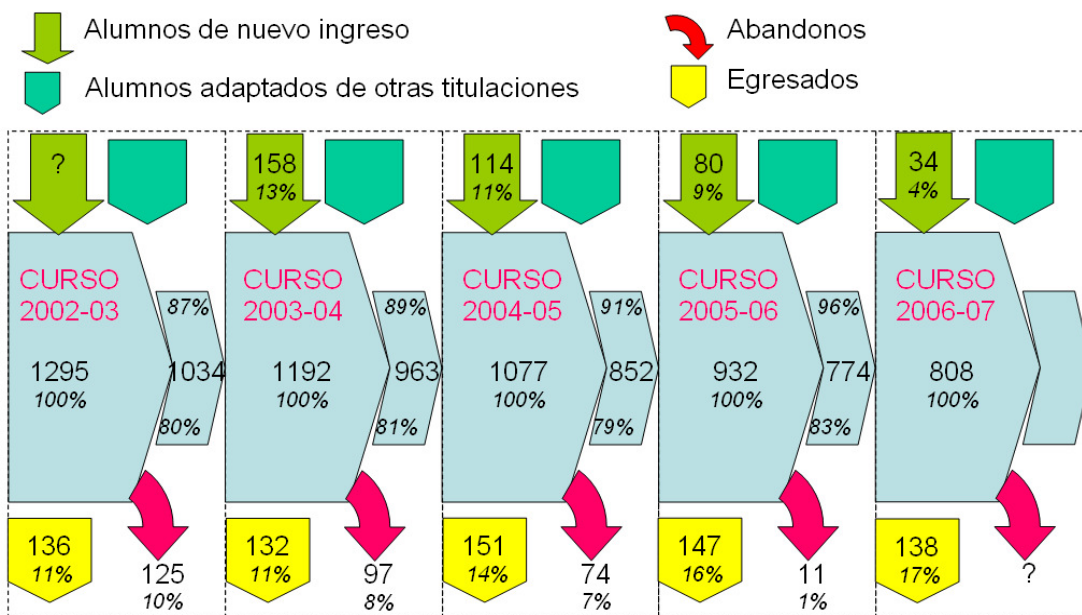


Figura 2. Diagrama de flujo de la titulación Ingeniero Informático en una de las universidades participantes

En este gráfico se representa, para cada curso, el total de alumnos matriculados en la titulación, que supone el 100 % para ese curso. Se representan, con las flechas respectivas, los alumnos de nuevo ingreso, los posibles adaptados de otras titulaciones, los egresados y los abandonos (alumnos que no son egresados y que no aparecen matriculados al curso siguiente); todos ellos junto al porcentaje que representan con respecto al total de ese curso. Entre cada dos cursos hay una flecha que representa los alumnos que siguen matriculados en la titulación de un curso para el siguiente; los dos porcentajes representan el porcentaje que esos alumnos suponen respecto de los dos cursos entre los que se encuentran.

Este tipo de gráficos debe interpretarse con cuidado, ya que puede haber regímenes transitorios en los flujos de alumnos, debido a un cambio de plan de estudios, o cualquier otra circunstancia.

Tomando en consideración aquellas universidades que muestren un diagrama de flujo estable, se pueden construir diagramas de flujo agregados para una determinada titulación, que dan idea de los flujos de alumnos en la titulación en el conjunto de varias universidades.

5. Resultados y conclusiones

Como ya se ha mencionado anteriormente, los resultados de este trabajo no son, a día de hoy, públicos, pero la metodología de trabajo propuesta permite obtener valores de referencia de distintos indicadores de rendimiento académico en un conjunto de centros que imparten una misma titulación.

6. Agradecimientos

Nuestro agradecimiento a todas las personas que, directamente o indirectamente, han participado en el proyecto EA2007-0152.

7. Referencias

- [1] Informe del MEC en <http://web.micinn.es>⇒Servicios (Estadísticas e Indicadores) ⇒Universidades⇒Datos Generales⇒Datos y Cifras del Sistema Universitario Español Curso 05-06 (Septiembre 2008).
- [2] <http://web.micinn.es>⇒ Servicios (Estadísticas e Indicadores)⇒Universidades⇒Indicadores ⇒Catálogo de Indicadores. (Septiembre 2008).
- [3] MÁ S, J; ALCOVER, R; BENLLOCH, J; BLES A, P; PONT, A; PRIETO, N; ROBLES, A; VALIENTE, J; ZÚNICA, L. "Una herramienta informática para un estudio multidimensional del rendimiento académico en la EUI de la UPV". *Libro de resúmenes del VII Congreso Universitario de Innovación Educativa en las enseñanzas técnicas (CUIE)*. Huelva, 1999. pp 2279-2290.
- [4] MÁ S, J; VALIENTE, J; ZÚNICA, L; ALCOVER, R; BENLLOCH, J; BLES A, P. "Estudio de la influencia sobre el rendimiento académico de la nota de acceso y procedencia (COU/FP) en la E.U. de Informática". *Libro de resúmenes de las VIII Jornadas sobre la Enseñanza Universitaria de la Informática (JENU I 2002)*. Cáceres, 2002. pp. 197-204.
- [5] ZÚNICA, L; BLES A, P; ALCOVER, R; MÁ S, J; VALIENTE, J. "Estudio del rendimiento académico de una asignatura con relación a asignaturas de cursos anteriores". *Libro de resúmenes de las IX Jornadas sobre la Enseñanza Universitaria de la Informática (JENU I 2003)*. Cádiz, 2003.
- [6] BONET, P; FERRANDO, M; MARTÍNEZ, B; FERNÁNDEZ, A; BONET, E; BLES A, P; MÁ S, J. "Modelo de almacén de datos para una Universidad Europea en Red". *International Conference on Network Universities and E-learning*. Tarragona, 2003.

- [7] ZÚNICA, L; ALCOVER, R; MÁS, J; VALIENTE, J; BENLLOCH, J.V; BLESA, P. "Relación entre el rendimiento de dos asignaturas de segundo curso y las asignaturas de primer curso en Ingenierías Técnicas de Informática de la Universidad Politécnica de Valencia". *Libro de resúmenes del I Simposio Nacional de Docencia en la Informática (SINDI 2005)*. Granada, 2005. pp 9-16.
- [8] HERNÁNDEZ, J; RAMÍREZ, M.J; FERRI, C. *Introducción a la Minería de Datos*. Ed. Prentice Hall, ISBN: 84-205-4091 9, 2004.
- [9] WITTEN, I.H; FRANK, E. *Data Mining: Practical Machine Learning Tools and Techniques* (Second Edition). Ed. Morgan Kaufmann, ISBN 0-12-088407-0, 2005.
- [10] <http://www.spss.com/clementine> (Septiembre 2008).
- [11] <http://www.dsic.upv.es/~cferrri/weka/> (Septiembre 2008).