



Recurso docente para la atención de la diversidad en el aula mediante la predicción de notas

Maria Salamó, Inmaculada Rodríguez, Maite López-Sánchez,
Anna Puig, Simone Balocco, Mariona Taulé

Universitat de Barcelona

Gran Via de les Corts Catalanes, 585, 08007 Barcelona

maria.salamo@ub.edu, inmarodriguez@ub.edu, maite_lopez@ub.edu,
annapuig@ub.edu, simone.balocco@ub.edu, mtaule@ub.edu

Resumen

Desde la implantación del Espacio Europeo de Educación Superior (EEES) en los diferentes grados, se ha puesto de manifiesto la necesidad de utilizar diversos mecanismos que permitan tratar la diversidad en el aula, evaluando automáticamente y proporcionando una retroalimentación rápida tanto al alumnado como al profesorado sobre la evolución de los alumnos en una asignatura. En este artículo se presenta la evaluación de la exactitud en las predicciones de GRADEFORSEER, un recurso docente para la predicción de notas basado en técnicas de aprendizaje automático que permite evaluar la evolución del alumnado y estimar su nota final al terminar el curso. Este recurso se ha complementado con una interfaz de usuario para el profesorado que puede ser usada en diferentes plataformas *software* (sistemas operativos) y en cualquier asignatura de un grado en la que se utilice evaluación continuada. Además de la descripción del recurso, este artículo presenta los resultados obtenidos al aplicar el sistema de predicción en cuatro asignaturas de disciplinas distintas: Programación I (PI), Diseño de Software (DSW) del grado de Ingeniería Informática, Tecnologías de la Información y la Comunicación (TIC) del grado de Lingüística y la asignatura Fundamentos de Tecnología (FDT) del grado de Información y Documentación, todas ellas impartidas en la Universidad de Barcelona. La capacidad predictiva se ha evaluado de forma binaria (aprueba o no) y según un criterio de rango (suspense, aprobado, notable o sobresaliente), obteniendo mejores predicciones en los resultados evaluados de forma binaria.

Palabras clave: Aprendizaje automático, Sistema de predicción de notas, Herramienta docente.

Recibido: 17 de noviembre de 2015; **Aceptado:** 26 de diciembre de 2015.

1. Introducción

La adaptación de los distintos grados al Espacio Europeo de Educación Superior (EEES) ha significado un cambio en el modelo de enseñanza universitaria donde el estudiante ha pasado a ser el sujeto activo del aprendizaje convirtiéndose en un generador de información, ya sea de datos de sus distintas actividades docentes, como de interacciones con plataformas de *e-learning*. La utilización de plataformas digitales para la interacción en el proceso de enseñanza-aprendizaje ha abierto la posibilidad de analizar esta información tanto para la reflexión sobre el propio proceso educativo, como para la predicción de la evolución del estudiante. Estos dos factores son cruciales para entender el enfoque del Análisis del Apre-

dizaje (*Learning Analytics*), que se define como la medida, recopilación, análisis y presentación de datos sobre el alumnado (y sus contextos), y que tiene la finalidad de comprender y optimizar el aprendizaje [14].

El Análisis del Aprendizaje abarca diferentes dimensiones que caracterizan el marco de actuación, como pueden ser los usuarios a los que va dirigido o los métodos utilizados [2]. En este artículo se presenta la exactitud en la predicción de notas de la plataforma GRADEFORSEER, un nuevo recurso docente de análisis de los resultados del aprendizaje para estimar la evolución del alumnado en una asignatura, facilitando al profesorado una predicción sobre el éxito o fracaso que pueda tener cada estudiante en una asignatura. A partir de los primeros resultados de pruebas, actividades, ejercicios

o prácticas de cada estudiante en una asignatura cuya evaluación se realiza de forma continuada a lo largo del semestre, GRADEFORSEER predice automáticamente el resultado final más probable del estudiante. El profesorado puede rectificar y personalizar la curva de aprendizaje del alumnado en etapas muy preliminares, reforzando los puntos débiles con enlaces de interés, referencias bibliográficas, apuntes, tutorías personalizadas, ejercicios resueltos, etc. En este sentido, a diferencia de los Sistemas Tutores Inteligentes [3], nuestra propuesta no sustituye al profesorado sino que detecta qué alumnado no sigue de forma satisfactoria los objetivos de aprendizaje para que el profesor pueda incidir en su planificación docente.

Para evaluar la precisión de las predicciones realizadas en la plataforma GRADEFORSEER durante el proceso de aprendizaje, en este artículo se presentan los resultados obtenidos en asignaturas de distintos grados de la Universidad de Barcelona. En primer lugar, se han analizado asignaturas de los primeros cursos del Grado de Ingeniería Informática, en los que el abandono y el fracaso es elevado. Concretamente, se ha analizado la precisión de las predicciones en las asignaturas de Programación I y Diseño de Software. Actualmente, el alumnado de estas asignaturas proviene de estudios de Bachillerato y de módulos de Grado Superior, con unas dinámicas de trabajo y retroalimentación de sus actividades muy diferentes a las implantadas en la universidad. En estos estudios previos, la ratio de alumnos por aula es muy reducida, los alumnos tienen numerosas pruebas, una o varias entregas semanales de pequeños ejercicios y una planificación muy guiada de los trabajos que deben realizar. Al llegar a la universidad, la ratio de alumnos por aula es superior, las pruebas están más dispersas en el tiempo, las entregas de ejercicios prácticos son a medio plazo y se promueve que cada estudiante planifique su trabajo de forma autónoma. Este cambio de dinámica es un factor esencial que influye en la pérdida de motivación y el consecuente abandono de estudiantes en las asignaturas de los primeros cursos.

Cabe destacar que el sistema es fácilmente generalizable a otras asignaturas del grado, siempre que el plan docente de estas asignaturas contemple la realización de diversas pruebas de evaluación (prácticas, parciales, presentaciones orales, etc.) a lo largo del curso. Así pues, en este artículo se han incluido también los resultados obtenidos en la predicción automática de las notas de las asignaturas de Tecnologías de la Información y la Comunicación del grado de Lingüística y la asignatura Fundamentos de Tecnología del grado de Información y Documentación, impartidas en la Facultad de Filología y la Facultad de Biblioteconomía y Documentación de la Universidad de Barcelona.

Para facilitar la utilización del sistema y la visualización de la predicción, el sistema se ha complementado con una aplicación multiplataforma dirigida al profesorado que puede ser utilizada en cualquier contexto (diferentes asignaturas).

2. Trabajo relacionado

Dentro del ámbito docente, el abanico de posibilidades que se abre a la hora de introducir herramientas *software* para hacer un seguimiento del alumnado es realmente amplio y diverso [6, 3, 15]. En este apartado no se pretende cubrir de forma exhaustiva dicha variedad, sino mencionar algunos de los trabajos que son relevantes para contextualizar y diferenciar la aportación del recurso docente que aquí se presenta.

Los sistemas más populares dedicados al seguimiento del alumnado son los Sistemas Tutores, que al incluir técnicas del área de Inteligencia Artificial son denominados Sistemas Tutores Inteligentes [3, 11]. En general, estos sistemas pretenden sustituir el tutor (entendido como profesor, alguien que guía al alumnado en el aprendizaje de una materia), de manera que es el propio sistema el que marca la forma en la que el alumnado adquiere nuevas competencias y conocimientos y que evalúa si éste alcanza los objetivos de aprendizaje marcados. Sin embargo, en el caso de nuestro recurso docente GRADEFORSEER, es importante destacar que su objetivo es ayudar al profesorado en el seguimiento de la evolución del alumnado proporcionándole indicadores de su rendimiento académico, y alertándole cuando se prevee que dicho rendimiento no es el esperado para que el profesorado pueda tomar medidas preventivas. Por tanto, este recurso no sustituye en ningún caso al profesorado, ya que requiere que éste siga llevando a cabo sus tareas docentes, tanto de acompañamiento en la adquisición de competencias y conocimiento como de evaluación.

Siguiendo esta misma línea de discurso, es importante diferenciar el recurso docente propuesto en este artículo del aprendizaje electrónico (*e-learning*) [15], un área que fusiona las Tecnologías de la Información y Comunicación (TIC) y la educación para permitir, en primera instancia, el aprendizaje autónomo (autoaprendizaje) a distancia. Como no podía ser de otro modo, los Sistemas Tutores constituyen componentes fundamentales en los sistemas de aprendizaje electrónico [10], y de la misma forma que hablamos de Sistemas Tutores Inteligentes, es posible aplicar técnicas de Inteligencia Artificial (como por ejemplo, la personalización) en los sistemas de aprendizaje electrónico [7].

Sin embargo, si nos centramos en el seguimiento del alumnado sin incidir en las tareas típicas del profesor (enseñanza y evaluación), encontramos que el concepto de tutor se entiende también como de orientador individualizado, alguien que pueda detectar problemas y disparar las alarmas necesarias para la reorientación del alumnado [1]. Esta es la aproximación que toma nuestro recurso docente y que se localiza dentro del área del Análisis del Aprendizaje [14]. Este enfoque se puede abordar realizando el seguimiento dentro de una asignatura concreta (tal y como es en nuestro caso¹) o, de forma transversal a los estudios que se estén realizando. Esta última aproximación se está empezando a implantar en algunas universidades [5], aunque su implantación es aún discreta, ya que el *Education Advisory Board* de Estados Unidos calcula

¹El módulo de aprendizaje propuesto ha sido desarrollado dentro de un proyecto de innovación docente [12].

que únicamente unas 125 instituciones universitarias del total de las 4000 (aprox.) existentes aplican técnicas de análisis del aprendizaje para seguir la evolución del alumnado a lo largo de los estudios. No obstante, ya se han aportado resultados positivos respecto a la reducción de alumnado que abandona los estudios [8].

3. GRADEFORSEER

En esta sección vamos a detallar cómo se ha abordado el recurso docente desde la perspectiva de un problema de predicción y se ha resuelto mediante diferentes técnicas de aprendizaje automático definidas en un módulo de predicción independiente. Además, se introduce la interfaz de usuario GRADEFORSEER-I, que permite al profesorado solicitar y obtener una predicción de notas de los alumnos de sus asignaturas.

3.1. Sistema de predicción

Tal y como se muestra en la figura 1, la idea principal del módulo de predicción es utilizar los datos históricos de una asignatura que contengan los resultados del alumnado a lo largo de uno o más cursos académicos y, mediante diferentes algoritmos de clasificación,² predecir la nota final que obtendrá el alumnado del curso académico actual.

El objetivo es por tanto predecir la nota de los alumnos en un instante del curso académico en el que ya se han realizado diversas³ pruebas de cualquier tipo que evalúen competencias básicas o transversales, ya sean prácticas, parciales o presentaciones en clase, pero que tengan asignada una calificación. De esta forma, por ejemplo a mitad de curso, el profesorado puede analizar la evolución de sus alumnos y así realizar las acciones necesarias para reforzar los puntos más débiles detectados en aquellos alumnos susceptibles de suspender la asignatura. Asimismo, la predicción de notas puede ser una herramienta para motivar a alumnos que, con las notas que tienen actualmente, ven difícil aprobar. De esta forma, se aborda la frustración que sufren los alumnos que obtienen calificaciones bajas en las primeras pruebas del curso.

Para el desarrollo de GRADEFORSEER se ha utilizado Java, porque tiene la ventaja que es multiplataforma y se espera que el recurso docente sea utilizado por diferentes usuarios en diferentes plataformas *software*. Asimismo, las librerías de aprendizaje automático que incorporan los diferentes clasificadores también están implementadas en Java, lo que nos ha permitido el desarrollo de un módulo homogéneo.

En concreto, tal y como se puede observar en el diagrama conceptual de la figura 1, el módulo de predicción realiza tres pasos: (1) recoge los datos de entrada; (2) entrena un conjunto de clasificadores con los datos recibidos; y (3) genera las predicciones y las vuelca en un fichero de salida.

En el primer paso, el módulo de predicción recoge como datos de entrada tanto los datos históricos completos de una asignatura, los cuales serán el conjunto de entrenamiento de los diferentes clasificadores, como los datos incompletos de los alumnos actuales, que corresponderán con el conjunto a predecir por cada clasificador. El histórico de cada curso académico (de cada asignatura) contiene un conjunto de alumnos y para cada uno de ellos, el conjunto de notas obtenidas en los ejercicios, prácticas, pruebas o exámenes, así como la nota final obtenida en dicha asignatura al finalizar el curso.

En el segundo paso, a partir de estos datos (el histórico y los datos a predecir) el módulo de predicción ejecuta diversos algoritmos de clasificación. Estos provienen de dos librerías de aprendizaje automático que se han unificado en el módulo: CBR-KM [9] y WEKA [4], cada una de ellas con un conjunto de algoritmos de clasificación.

- CBR-KM (*Case-Based Reasoning-Knowledge Management*) es una librería Java de aprendizaje automático desarrollada en la Universidad de Barcelona que incorpora un conjunto de algoritmos de clasificación de Razonamiento Basado en Casos y algoritmos específicos para el mantenimiento de la memoria de casos, entre ellos el *Adaptive Case-Based Reasoning* (ACBR) [13].
- WEKA (*Waikato Environment for Knowledge Analysis*) es una librería Java de aprendizaje automático desarrollada por la Universidad de Waikato que es de *software* libre bajo licencia GNU. Esta librería contiene algoritmos para el preprocesado de los datos, algoritmos de clasificación y de clusterización, diferentes métodos de evaluación, además de una interfaz para visualizar datos y comparar algoritmos. En el desarrollo de nuestro módulo de predicción se han usado tanto los algoritmos para el preprocesado de los datos, como los algoritmos de clasificación disponibles en la librería. Esta librería incorpora un total de 87 algoritmos de clasificación provenientes de WEKA como por ejemplo algoritmos de redes bayesianas, redes neuronales, árboles de decisión, algoritmos perezosos, etc.

En el tercer paso, el módulo de predicción extrae un fichero de datos con la nota final estimada para cada alumno del curso actual. Actualmente esta nota se predice haciendo una validación cruzada (ver más detalles en la sección 4.2.1) del histórico (conjunto de entrenamiento) y, a partir del mejor clasificador obtenido en esta validación cruzada, se procede a predecir la nota final de los alumnos del curso actual.

Finalmente, cabe mencionar que la idea de usar un sistema clasificador para predecir la evolución de un alumno ya se planteó en un trabajo anterior [16]. En concreto, se analizó como un clasificador de Razonamiento Basado en Casos predecía la evolución de los alumnos en la asignatura de Programación I impartida en la Universidad Ramon Llull. A di-

²Un algoritmo de clasificación tiene como objetivo asignar un elemento entrante no etiquetado en una categoría concreta conocida.

³Aunque el módulo de predicción funcionará mejor con un mayor número de notas, las evaluaciones realizadas demuestran que un 50 % del total es suficiente para obtener resultados aceptables.

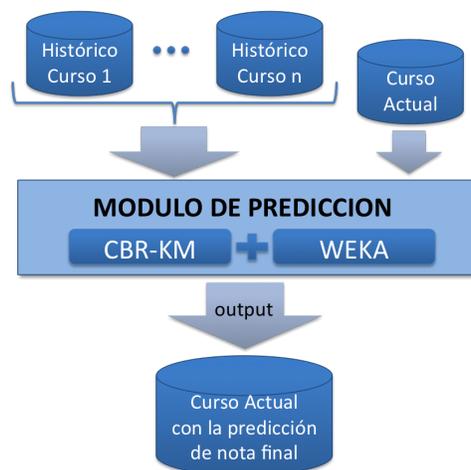


Figura 1: Diagrama conceptual del sistema predictivo.

ferencia del trabajo anterior, en el que se analizó un único clasificador y se desarrolló para una asignatura en concreto, este artículo presenta un recurso docente desarrollado en el contexto de un proyecto de innovación docente [12] que se adapta a cualquier asignatura, que engloba un módulo con un conjunto de clasificadores provenientes de dos librerías para realizar la predicción y que incluye, además, una interfaz que facilita el uso al profesorado.⁴

3.2. Interfaz GRADEFORSEER-I

En esta sección se presenta GRADEFORSEER-I, una aplicación para predecir la nota final del alumnado de una asignatura. Se trata de un prototipo que tiene como usuarios principales a profesores de informática, aunque también está abierto a toda la comunidad educativa y por este motivo su interfaz se ha adaptado a un perfil de usuario menos técnico.

El diseño de la interfaz de usuario sigue el patrón denominado Wizard que permite al usuario realizar su tarea en una serie de pasos bien definidos. Focalizada en principios de usabilidad tales como la aprendibilidad y memorabilidad, esta interfaz se adecúa tanto a perfiles de usuarios no expertos en la tarea a realizar como a aquellos usuarios que la realizan de forma esporádica.

GRADEFORSEER-I es una aplicación de escritorio multiplataforma, desarrollada en Java, que utiliza Java Swing para la interfaz gráfica de usuario (Graphical User Interface, GUI). Durante la instalación de la aplicación se permite seleccionar el idioma de la interfaz y se instala automáticamente la librería de clasificadores de WEKA y CBR-KM, necesaria para su correcta ejecución.

Para que la aplicación pueda realizar la predicción de notas de los alumnos de una asignatura, esta necesita primero que el profesor proporcione i) las notas de las diferentes actividades evaluativas de la asignatura en cursos anteriores (los

datos históricos) y ii) las notas que tiene de dicha asignatura en el curso actual. El objetivo es predecir a partir de los datos (incompletos) de que se dispone hasta el momento del curso actual (por ejemplo, las 3 primeras prácticas y la primera prueba parcial teórica), cuál podría ser la nota final de los alumnos.

Estos datos se deben proporcionar en los formatos CSV o XLS, en un fichero con nombre <NombreAsignatura>_Curso_<AñoInicial>_<AñoFinal>.xls. Por ejemplo, el fichero de la asignatura de Diseño de Software del curso 2013-2014 se llamará «DSW_Curso_2013_2014.xls». Tal y como se ha mencionado, estos ficheros deben contener para cada alumno la nota obtenida en cada una de las actividades evaluativas (prácticas, pruebas parciales, finales).

Una vez instalada la aplicación, y suponiendo que el profesor tiene preparados los ficheros CSV o XLS mencionados previamente, puede comenzar a interactuar con la interfaz realizando los siguientes cuatro pasos:

1. Cargar los datos históricos: uno o más ficheros con las notas de cursos académicos anteriores.
2. Cargar los datos actuales: un fichero con las notas del curso actual.
3. Solicitar la predicción: el profesor puede seleccionar la opción *Predicción rápida* o bien *Predicción exhaustiva* (más lenta).
4. Visualizar el resultado de la predicción.

En el primer paso (figura 2) el usuario proporciona las notas de cursos académicos anteriores. En este paso, también se puede visualizar el contenido de los ficheros proporcionados previamente o eliminar algún fichero que finalmente no se quiera incluir en el histórico con el que trabajará el algoritmo de predicción. En el segundo paso (figura 3), el usuario proporciona los datos del curso actual.

⁴El código del recurso docente se encuentra disponible en <http://wai.maia.ub.es/projects.htm> (Educational innovation).

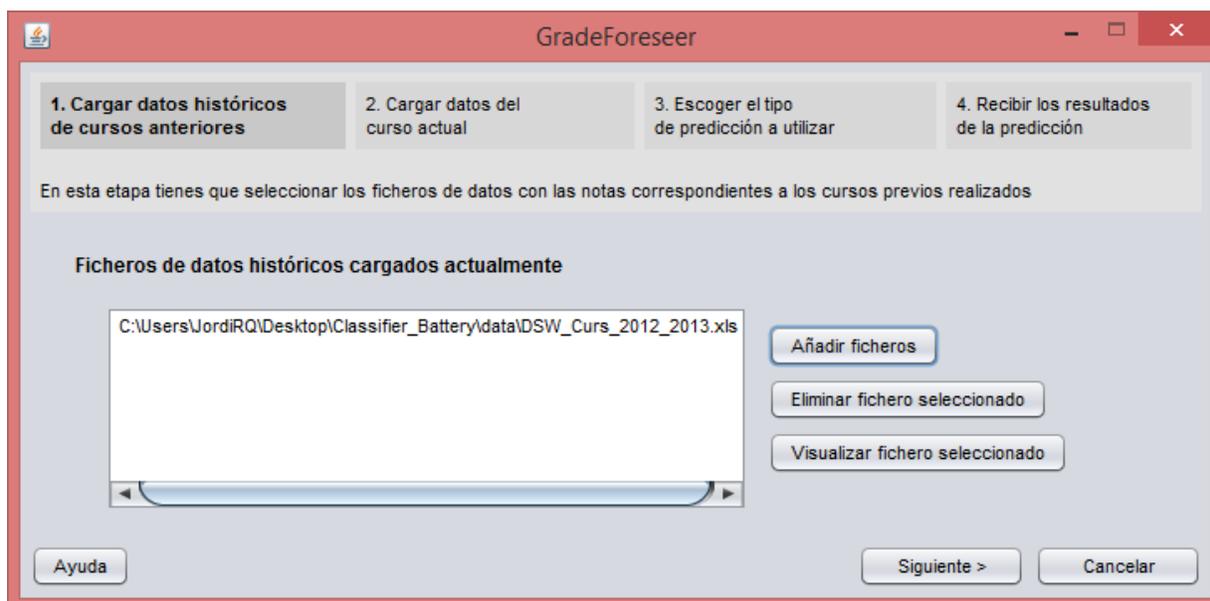


Figura 2: GRADEFORSEER-I Carga de los datos históricos.

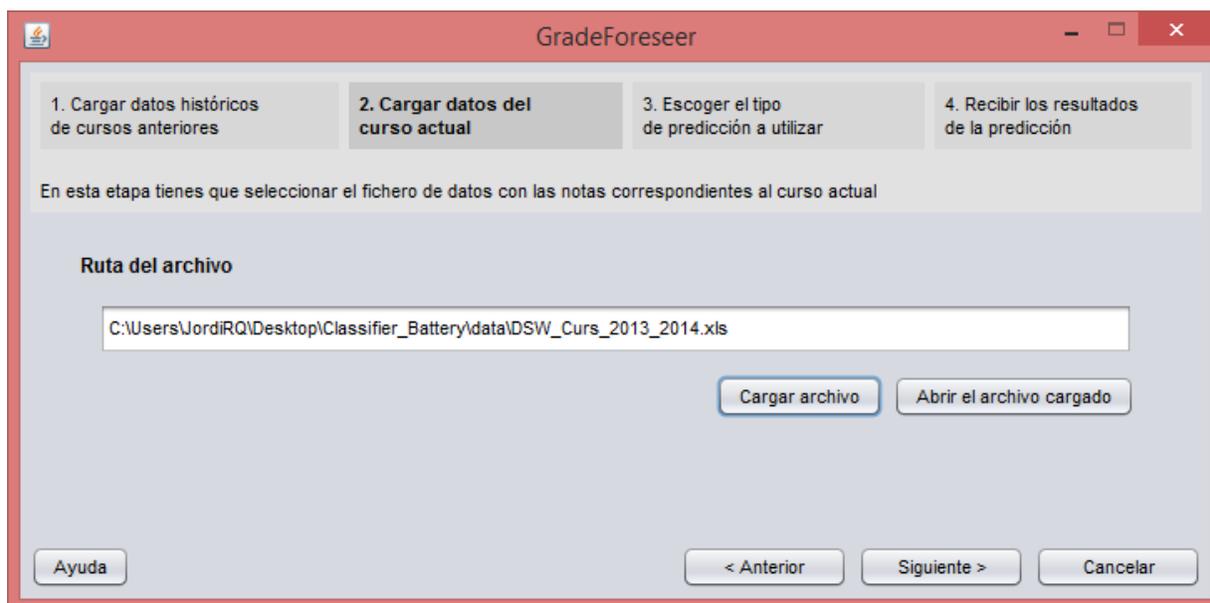


Figura 3: GRADEFORSEER-I Carga de los datos actuales.

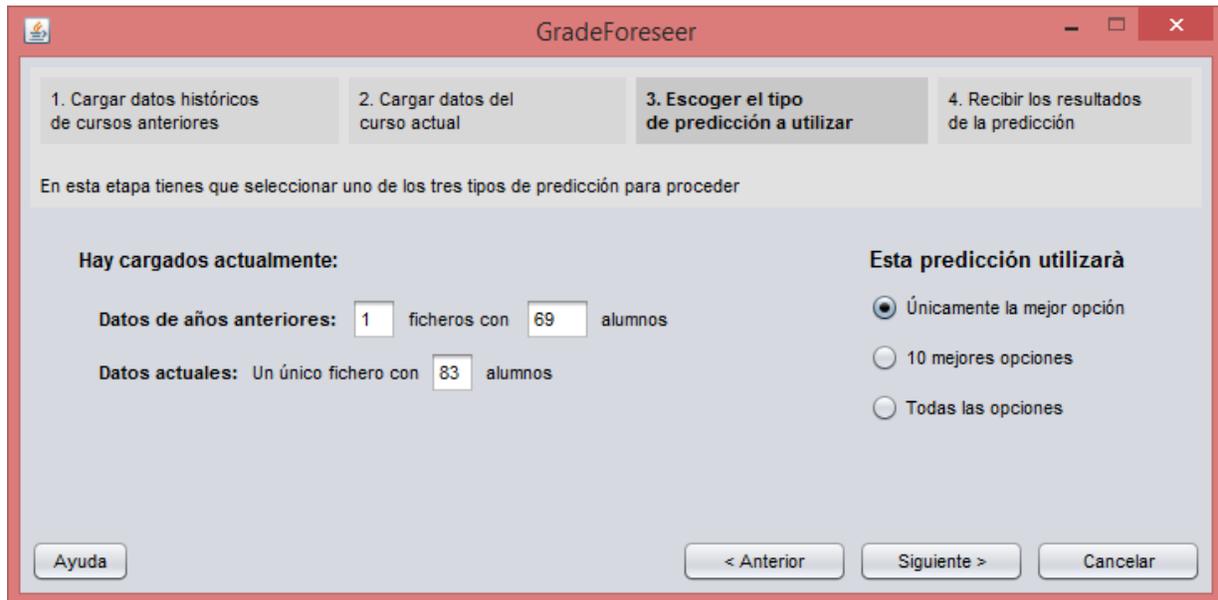


Figura 4: GRADEFORSEER-I Selección del tipo de predicción.

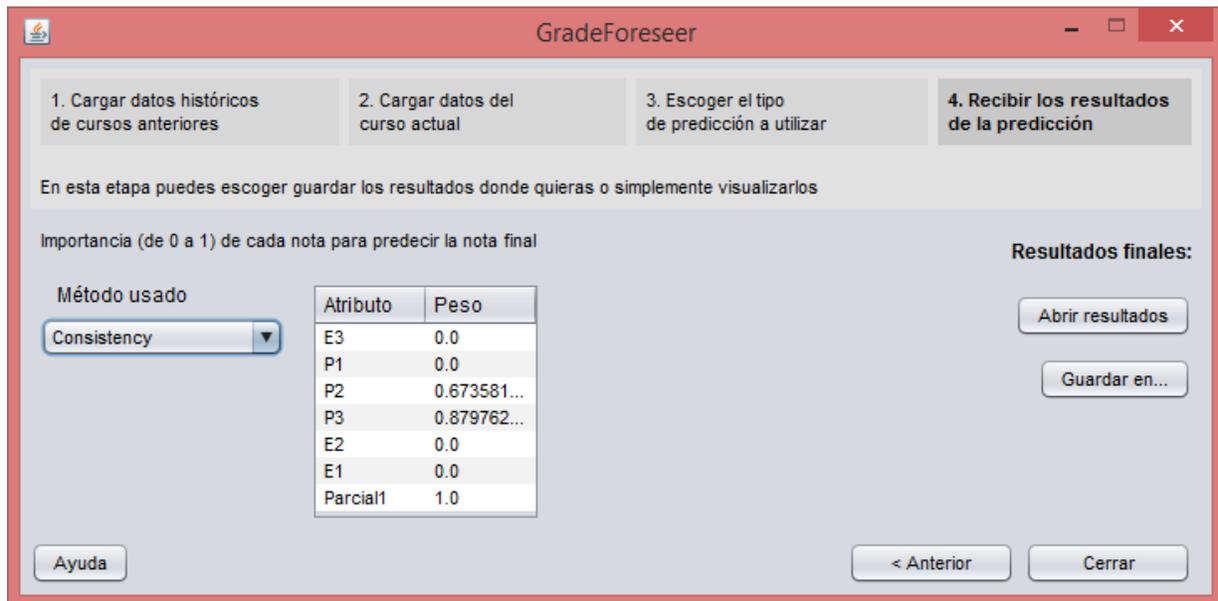


Figura 5: GRADEFORSEER-I Visualización de los resultados.

Una vez cargados los ficheros de notas (históricos y actual), en el tercer paso el profesor está en disposición de solicitar la predicción (figura 4) a partir de tres opciones en función del número de algoritmos de predicción que se quieran aplicar. La primera opción, «Únicamente la mejor opción», es una predicción rápida que utiliza un único algoritmo de predicción de las librerías presentadas en la sección 3.1. La segunda opción, «10 mejores opciones», es una predicción exhaustiva que realiza pruebas de validación cruzada con el conjunto de datos de entrenamiento (datos de cursos anteriores) con los 10 mejores clasificadores de toda la batería de clasificadores incluidos en el módulo de predicción. Estos clasificadores han sido elegidos a partir de un análisis exhaustivo con validación cruzada en diferentes asignaturas y, en concreto, se utilizan los 10 que en promedio obtuvieron mejores resultados. Esta segunda opción es, por lo tanto, más lenta que la anterior pero podría dar resultados más precisos. Finalmente, la tercera opción, «Todas las opciones», la más lenta de las tres, analiza los datos con todos los clasificadores del módulo de predicción y muestra el mejor resultado obtenido.

El último paso (figura 5) consiste en la visualización de los resultados una vez terminada la ejecución de los algoritmos de predicción. En concreto, incorpora también una opción para abrir el fichero que contiene la predicción de notas. Este fichero contiene una nueva columna con la nota prevista: Suspenso, Aprobado, Notable o Sobresaliente. Finalmente, el profesor puede guardar los datos de predicción generados por la aplicación si así lo desea. Además, en este último paso se puede observar la importancia o relevancia que tiene cada actividad evaluativa en la predicción de la nota final. Esta relevancia se calcula mediante métodos de ponderación de atributos, que permiten ver qué actividades evaluativas (ejercicios, prueba o actividad) son los más determinantes para predecir la nota final de un alumno. De este modo, el profesorado puede analizar los resultados concretos de un ejercicio, prueba o actividad con mayor relevancia (el máximo está en 1) y personalizar y reforzar las acciones para mejorar el aprendizaje de los temas tratados en esas actividades evaluativas.

4. Evaluación

La evaluación del sistema de predicción de notas se ha realizado con cuatro asignaturas impartidas en tres grados de la Universidad de Barcelona. Los grados se desarrollan a lo largo de 4 cursos académicos. Las asignaturas de Programación I (formación básica) y Diseño de Software (obligatoria), son asignaturas de primer y segundo curso respectivamente del grado de Informática. La asignatura de Fundamentos de Tecnología (formación básica) se imparte en el grado de Información y Documentación. Finalmente, la asignatura de Tecnologías de la Información y la Comunicación (obligatoria) se imparte en el grado de Lingüística.

4.1. Descripción de las asignaturas

La asignatura de Diseño de Software (DSW) tiene como objetivo mejorar las capacidades de programación e introducir al alumno en los conceptos básicos del diseño de *software*. Esta asignatura se imparte a continuación de las asignaturas de formación básica (como por ejemplo Programación I) y, en el plan de estudios, es la primera asignatura dedicada a la Ingeniería del Software. Así pues, la asignatura de DSW representa un reto para los alumnos de informática dado que están a caballo entre afianzar sus conocimientos en programación y empezar a desarrollar problemas de mayor envergadura.

La asignatura de DSW se imparte en clases teóricas (2 horas semanales) y clases prácticas (2 horas semanales). Para las prácticas se sigue el aprendizaje basado en problemas, donde se proporciona a los alumnos un problema de tamaño medio que se va resolviendo siguiendo principalmente las etapas de desarrollo del *software* del proceso unificado. En esta asignatura se realiza una evaluación continuada de los alumnos, que consta de tres entregas de pequeños ejercicios prácticos (*E1*, *E2*, *E3*), tres entregas de prácticas (*P1*, *P2*, *P3*) y dos exámenes (uno *Parcial1* a mitad del curso y otro *Parcial2* al acabar el curso). Por lo tanto, el histórico de notas de la asignatura consta de: 3 notas de ejercicios, 3 notas prácticas, 2 notas de pruebas teóricas y la nota final.

Programación I (PI) tiene como objetivo iniciar al estudiante en el mundo de la programación para resolver problemas científicos y técnicos de forma metódica y sistemática. La asignatura se imparte en clases teóricas (2 horas semanales) y clases prácticas (2 horas semanales). Las clases teóricas se dedican a la exposición del temario y a realizar ejercicios en pizarra. Las clases prácticas se realizan en aulas de informática donde los alumnos realizan 4 trabajos prácticos (*P1*, *P2*, *P3*, *P4*) a lo largo del curso. Las actividades de evaluación que se realizan en esta asignatura constan de una nota de teoría y una nota de prácticas. La nota de teoría se calcula con el resultado obtenido en dos pruebas parciales teóricas a lo largo del curso (*Parcial1*, *Parcial2*) y la nota de prácticas se calcula con los resultados obtenidos en cada una de las prácticas entregadas. Por lo tanto, el histórico de notas que se puede proporcionar al sistema de predicción consta de: 4 notas de prácticas, 2 notas de pruebas teóricas y la nota final.

Fundamentos de Tecnología (FDT) es una asignatura que se imparte en la facultad de Biblioteconomía a estudiantes que no tienen una preparación científica o informática previa. El objetivo de este grado no es el de enseñar a programar, sino el de proporcionar los conocimientos necesarios a futuros usuarios de la tecnología que la aplicaran a problemas concretos del mundo económico, de la gestión de la información y de la documentación. Se explican conceptos básicos de tecnología que permitan a los alumnos utilizar herramientas informáticas en su trabajo. La asignatura se imparte en clases teóricas (2 horas semanales) y clases prácticas (2 horas semanales). Las clases teóricas se dedican a la exposición del temario por medio de transparencias y esquemas en la pizarra. Las clases prácticas se realizan en aulas de informática donde los alum-

	A	B	C	D	E	G
1	ID_user	E1	E2	E3	Parcial1	NOTA PREVISTA
2	1	10	1	1	3,5	Aprobado
3	2	0	0	0	0	Suspenso
4	3	0	6	0	5	Aprobado
5	4	4	0	0	0	Suspenso
6	5	4	8,5	7,5	7	Aprobado
7	6	10	9	9	8	Excelente
8	7	0	0	0	3,75	Suspenso
9	8	9	0	0	2,75	Aprobado
10	9	10	9,8	9,75	6,55	Excelente
11	10	5	9,5	7,5	2,35	Aprobado
12	11	0	3	0	3,75	Suspenso
13	12	0	0	0	0	Suspenso
14	13	5	7	4,5	3,25	Suspenso
15	14	0	0	0	0	Suspenso
16	15	7,8	9,5	0	4	Suspenso
17	16	7	5	5,5	5	Suspenso
18	17	2	0	0	0	Suspenso
19	18	6,6	8	9	5,5	Notable

Figura 6: Fichero con una nueva columna que muestra el resultado de la predicción.

nos realizan 3 trabajos prácticos ($P1$, $P2$, $P3$) a lo largo del curso. Las actividades de evaluación que se realizan en esta asignatura constan de una nota de teoría y una nota de prácticas. La nota de teoría se calcula con el resultado obtenido en dos pruebas parciales teóricas a lo largo del curso ($Parcial1$, $Parcial2$) y la nota de prácticas se calcula con los resultados obtenidos en cada una de las prácticas entregadas. Por lo tanto, el histórico de notas que se puede proporcionar al sistema de predicción consta de: 3 notas prácticas ($P1$, $P2$ y $P3$) y 2 notas de pruebas teóricas ($Parcial1$, $Parcial2$).

Tecnologías de la Información y la Comunicación (TIC) es una asignatura obligatoria del grado de Lingüística de la Facultad de Filología que se imparte en el segundo curso. Se trata de una asignatura fundamentalmente práctica (90 % del curso) cuyo objetivo es proporcionar a los estudiantes herramientas básicas del ámbito de las TIC relacionadas con el tratamiento de textos escritos. Las sesiones prácticas se realizan en el aula de informática y se agrupan en tres tipos de actividades: 1) la utilización de programas para el análisis cuantitativo y cualitativo de textos; 2) el procesamiento de textos con Unix y 3) la realización de formularios y la gestión de los datos recogidos utilizando una hoja de cálculo. Los alumnos tienen que realizar también un trabajo escrito individual que consiste en plantear una propuesta de investigación sobre un corpus textual aplicando los distintos programas y herramientas presentados a lo largo del curso. La evaluación de la asignatura es continuada y consiste en la realización de un examen práctico (50 % de la nota final), un examen teórico (10 % de la nota final), la entrega de dos ejercicios prácticos (20 % de la nota final) y del trabajo (10 % de la nota final). Por lo tanto, el histórico de notas de la asignatura consta de: 2 notas de ejercicios ($E1$, $E2$), de un trabajo escrito ($E3$) y 2 de exámenes ($Parcial1$, $Parcial2$).

4.2. Resultados

Para identificar los clasificadores óptimos en cada asignatura y cuantificar el rendimiento de los mismos se han realizado dos experimentos numéricos: el primero evalúa los diferentes clasificadores con validación cruzada de datos de un curso académico y el segundo evalúa la capacidad predictiva de los mejores clasificadores en cada una de las asignaturas. Además, se ha realizado un tercer experimento con métodos de selección de atributos para detectar qué pruebas o ejercicios son los que más peso tienen en la determinación de la evolución de los alumnos en una asignatura.

En concreto, en cada uno de los experimentos se han entrenado los clasificadores usando los resultados de las asignaturas del penúltimo año y se han evaluado los resultados sobre las notas del último año académico. Para cada experimento se han usado las cuatro bases de datos disponibles: Diseño de Software, Programación I, Fundamentos de Tecnología y Tecnologías de la Información y la Comunicación. En total, se han ejecutado 87 clasificadores disponibles en las dos librerías CBR-KM y WEKA para cada una de estas bases de datos.

4.2.1. Experimento 1: Evaluación de los clasificadores

El objetivo del primer experimento es analizar qué clasificador es el que mejor predice los resultados en cada asignatura.

Para evaluar los resultados con un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y comprobación, se ha utilizado la validación cruzada de 10 iteraciones (*10 fold cross-validation*). Esta técnica consiste en utilizar una décima parte del conjunto de datos de entrada como datos de prueba y el resto (el 90 %) como datos de entrenamiento. El proceso de validación cruzada es

repetido durante 10 iteraciones, con cada uno de los posibles subconjuntos de datos de comprobación. En cada iteración se calcula el valor de exactitud que corresponde al número de aciertos sobre el número total de alumnos. Finalmente, se realiza la media aritmética de los valores de exactitud obtenidos en cada iteración para obtener un único resultado.

Además, se ha evaluado separadamente la capacidad del clasificador para predecir la nota final del alumno siguiendo dos criterios: por *rango* y *binario*. En el primer caso (*rango*), se han definido rangos de valores numéricos de notas $\{[1, 5]; [5, 7]; [7, 9]; [9, 10]\}$ correspondientes a las calificaciones de $\{Suspendido, Aprobado, Notable, Sobresaliente\}$, respectivamente. En el segundo (criterio *binario*), sólo se ha considerado si el estudiante aprueba o no la asignatura, es decir si su nota final es superior o inferior a 5.

En el cuadro 1 se han indicado los acrónimos de los clasificadores que han obtenido mejores resultados.

La figura 7 muestra para la asignatura de DSW el porcentaje promedio de la exactitud obtenida en los cinco mejores clasificadores. Tal y como se puede observar la exactitud es muy alta (entre el 87 % y el 91,6 %) en el caso de la clasificación binaria y cercana al 70 % en la clasificación por rangos.

En la figura 8, para la misma asignatura, se puede apreciar que la precisión y la sensibilidad para predecir por rango binario. En el caso del número de aprobados la precisión es muy buena, en promedio sobre el 90 %, siendo superior en la mayoría de casos. La precisión de los algoritmos de aprendizaje es, por lo tanto, muy buena para los aprobados y, en cambio, hay disparidad de resultados para los suspendidos. Por otro lado, la sensibilidad es superior al 95 % para los aprobados y moderada para los suspensos, alrededor del 60 %. Es decir, la exactitud es más difícil en el caso de alumnos que suspenden y que durante el curso su evolución tiende hacia el aprobado. Esto es debido principalmente al carácter incremental de la asignatura de DSW, siendo las últimas prácticas las más difíciles y las que ponderan más en la nota final del alumno.

En la figura 9 se muestran los resultados para la asignatura de PI. En esta asignatura también se observan resultados en la exactitud por encima del 90 % en la predicción binaria y alrededor del 80 % en la predicción por rangos. Por otro lado, la figura 10 muestra la precisión y sensibilidad obtenida en la clasificación por rango binario. En cuanto a la precisión, se observa que ésta puede llegar a ser del 97 % para los aprobados y del 89 % para los suspensos. Sin embargo, la sensibilidad en este caso no es tan marcada como en el caso de DSW y todos los algoritmos se mantienen en sensibilidades superiores al 80 %.

Tal y como muestra la figura 11, en la asignatura de FDT se puede apreciar que la exactitud está entre el 87 % y el 91 % en la predicción binaria y alrededor del 85 % para la predicción por rangos. En la figura 12, se observa que la precisión en promedio es aproximadamente del 88 % para los aprobados y del 94 % para los suspendidos. Sin embargo, la sensibilidad en este caso también es uniforme en aprobados y suspendidos, siendo los resultados superiores al 89 % para los tres mejores clasificadores.

Finalmente, la figura 13 y 14 muestran los resultados para la asignatura de TIC. En esta asignatura se obtiene una exactitud (véase figura 13) que está entre el 74 % y el 77 % para la predicción por rangos y entre el 86 % y el 91 % en la predicción por rango binario. En la figura 14 se observa que la precisión en el mejor caso es para el algoritmo BN, obteniendo un 90 % de precisión en los aprobados y un 91 % en los suspendidos. Sin embargo, la sensibilidad en este caso no es tan uniforme en aprobados y suspendidos, siendo más difícil de predecir los alumnos aprobados que los suspendidos. Esta asignatura también tiene un carácter incremental y las últimas evaluaciones son las que determinan la nota final de la asignatura.

En resumen, los resultados de la evaluación de la predicción de notas de los alumnos en las cuatro asignaturas demuestran que los algoritmos de clasificación pueden predecir cuidadosamente la nota final de un alumno, siendo en la mayoría de los casos la predicción superior al 85 %.

4.2.2. Experimento 2: Evaluación de la capacidad predictiva del clasificador

El interés del sistema predictivo es poder analizar los datos disponibles a lo largo de la asignatura y predecir la nota final del estudiante antes de disponer de todas las notas en un curso académico. Por este motivo, el objetivo del segundo experimento es analizar la capacidad predictiva del sistema en el caso que únicamente una parte de los resultados estén disponibles.

Para cada asignatura considerada, el conjunto de comprobación ha sido creado a partir de las notas del último año académico, omitiendo las últimas columnas de notas del curso. De este modo se han eliminado el 20 % en las asignaturas de FDT y TIC y el 33 % de los datos disponibles en DSW y PI. Esta disparidad de valores es debida al número total de pruebas o ejercicios de cada asignatura. En el caso de FDT y TIC, la eliminación de una prueba o ejercicio más significaba una reducción del 40 % de los datos. A partir de estos datos el sistema produce una predicción de nota final de la asignatura que se compara con las notas finales disponibles antes de la eliminación de datos.

En el cuadro 2, se presentan los tres mejores resultados de clasificación obtenidos en las cuatro asignaturas, ordenados de forma descendente desde el mejor resultado hasta el peor. En este caso, sólo se ha evaluado la capacidad del clasificador para predecir la nota final del alumnado siguiendo el criterio *binario*.

En este experimento los resultados de predicción son aceptables, aunque la exactitud obtenida en la predicción es ligeramente inferior a la propia del clasificador. Cuando el sistema de predicción evalúa los resultados de forma binaria, se obtiene un 88,0 % de exactitud en caso de la asignatura DSW, un 85,9 % para la asignatura PI, un 88,7 % para la de FDT y un 86,8 % para la de TIC.

Nombre del clasificador	Acrónimo
Averaged N-Dependence Estimator 1	A1DE
Averaged N-Dependence Estimator 2	A2DE
Bayes Network	BN
CBR Classifier	CBR
Dagging – Decision Stump	DDS
First Nearest-Neighbour	FNN
K-nearest-neighbor	KNN
Multilayer Perceptron	MP
Naive Bayes	NB
Naive Bayes Decision Tree	NBT
Ordinal Class Classifier	OCC
Rotation Forest	RF
Randomizable Filtered Classifier KNN	RFC
Sequential Minimal Optimization	SMO
Voting Feature Intervals	VFI

Cuadro 1: Acrónimo de los clasificadores utilizados en las figuras 7-14.

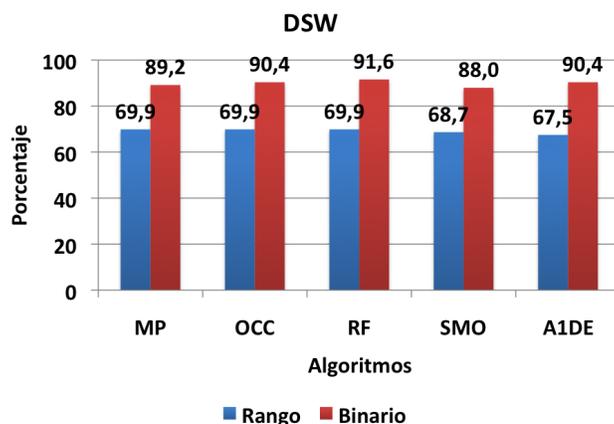


Figura 7: Exactitud en la asignatura de DSW.

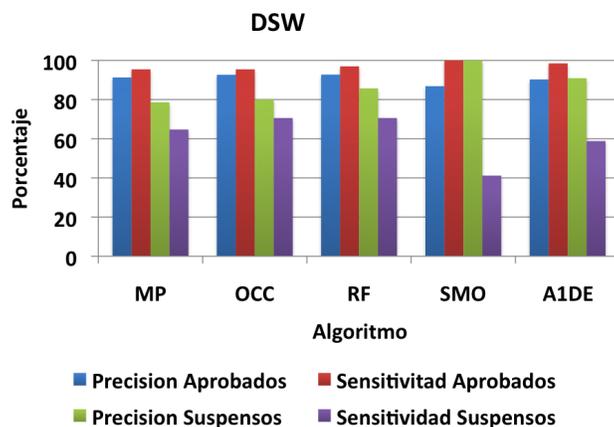


Figura 8: Precisión y sensibilidad en la asignatura de DSW.

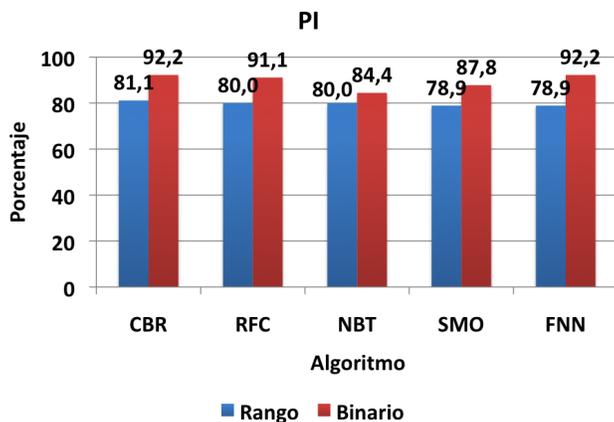


Figura 9: Exactitud en la asignatura de PI.

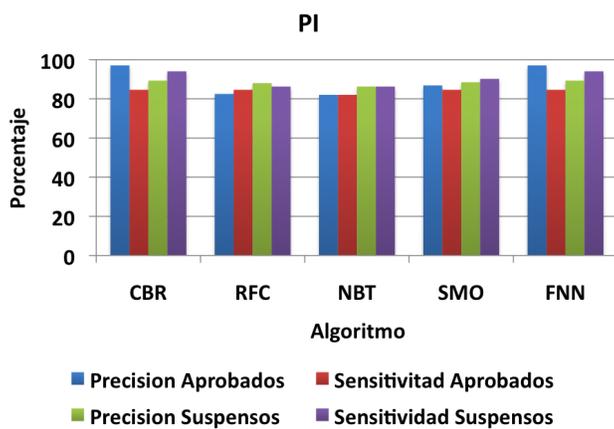


Figura 10: Precisión y sensibilidad en la asignatura de PI.

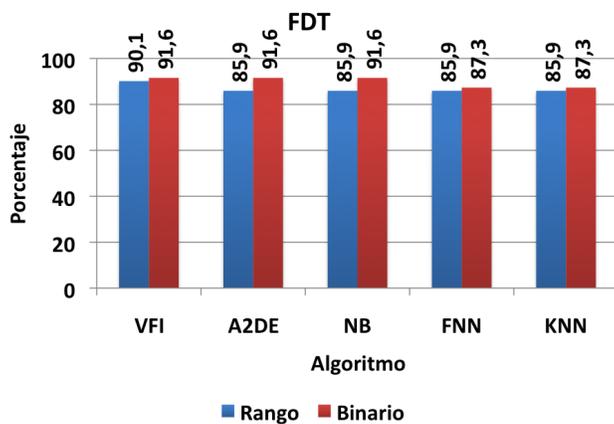


Figura 11: Exactitud en la asignatura de FDT.

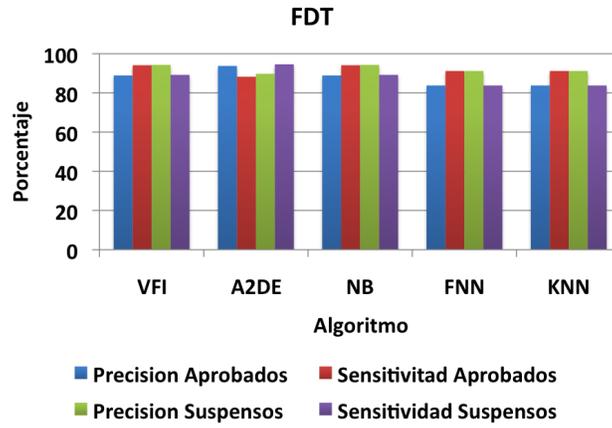


Figura 12: Precisión y sensibilidad en la asignatura de FDT.

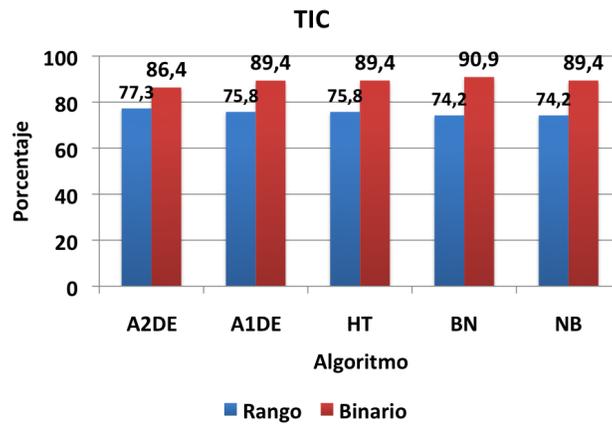


Figura 13: Exactitud en la asignatura de TIC.

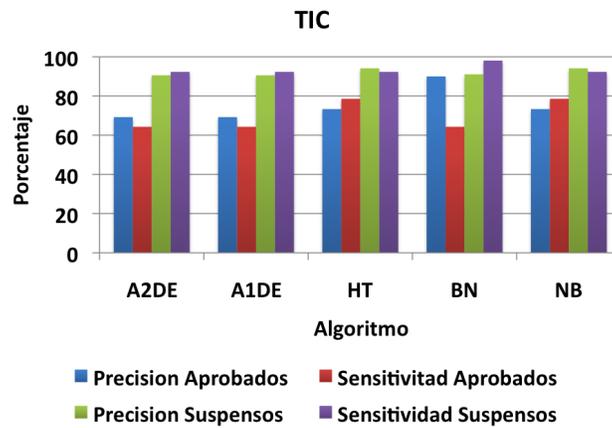


Figura 14: Precisión y sensibilidad en la asignatura de TIC.

Mejor clasificador	DSW	PI	FDT	TIC
#1	MP	BN	SMO	RF
Exactitud	87.95	85.87	88.73	86.79
#2	RF	A2DE	BN	A1DE
Exactitud	86.96	85.87	88.73	84.90
#3	MPCO	DDS	OCC	SMO
Exactitud	86.96	85.87	80.28	84.90

Cuadro 2: Resultados de los tres mejores clasificadores evaluados sobre cada asignatura con datos parciales para realizar una predicción binaria.

4.3. Experimento 3: Influencia de cada prueba o ejercicio

Para saber cuáles son las pruebas que más influyen en el resultado del alumno, se ha hecho un análisis mediante técnicas de ponderación automáticas para definir unos pesos a los diferentes elementos de evaluación de estas asignaturas. En concreto se han analizado con los siguientes métodos de ponderación: InfoGain, ChiSquare, Correlation, ReliefF, GainRatio, Significance y Simmetrical. El análisis también se ha realizado por diferentes criterios: por rango y binario. La metodología ha sido la misma que en el experimento 1, es decir, se ha usado validación cruzada de 10 iteraciones.

El cuadro 3 contiene los pesos asignados por todos los métodos de ponderación a cada una de las asignaturas usando una clasificación por rangos. En la asignatura de DSW, se observa que *E1*, *E2* y *E3* tienen un peso ínfimo, incluso de 0. Las prácticas tienen más peso, siendo *P2* la más importante seguida de *P3*. El examen parcial también es importante. Estos pesos corresponden con la realidad observada por el profesor que indica que los puntos críticos de los alumnos son la segunda y la tercera práctica, así como el segundo examen parcial (*Parcial2*) (el cual obtiene un peso de 1 en todos los algoritmos de ponderación).

En la asignatura de PI, hay una práctica inicial, *P1*, que todos los métodos consideran que no influye en la nota final obtenida por el alumno y, por lo tanto, no sirve para predecir el resultado final. En esta asignatura son muy importantes las notas obtenidas en los parciales (*Parcial1* y *Parcial2*), seguidas de las notas obtenidas en *P2*, *P3* y *P4*.

En la asignatura de FDT el peso más importante recae en el segundo examen parcial (*Parcial2*) para la mayoría de métodos de ponderación, seguido del *Parcial1*, en cambio, hay prácticas que no tienen prácticamente peso en la evaluación, como es el caso de *P3*. Por lo tanto, en esta asignatura es muy importante la nota obtenida en los exámenes parciales para poder predecir correctamente el resultado de un alumno.

Finalmente, en la asignatura de TIC se puede observar que hay ejercicios como *E1* que no tienen influencia en la evolución de un alumno en la asignatura y, en cambio, *E2* es un ejercicio que es determinante para predecir la nota final.

Todos los resultados observados permiten determinar en cada asignatura cuáles son las pruebas que más influyen en la

evaluación del alumno y dónde se puede detectar si finalmente un alumno aprobará o no la asignatura.

Se presentan en el cuadro 4 los resultados de qué elemento de la evaluación es más determinante en cada asignatura cuando se realiza una evaluación de los alumnos con un criterio binario.

En la asignatura de DSW los resultados obtenidos aplicando una evaluación según el criterio binario son muy similares al criterio por rangos. Se observa que *E1*, *E2* y *E3* tienen un peso ínfimo, incluso de 0. Las prácticas que tienen más peso son *P2* y *P3* y que el examen *Parcial2* es la más importante de todas las pruebas.

Por otro lado, en la asignatura de PI, los diferentes métodos de ponderación dan más importancia al *Parcial2*, *Parcial1* y a la última práctica del curso (*P4*). Al igual que en la evaluación por rangos, la práctica *P1* no se considera relevante para determinar si un alumno aprobará o no la asignatura.

En la asignatura de FDT se puede observar que según el método de ponderación hay disparidad de pesos para la evaluación binaria. Hay métodos de ponderación que consideran que *P2* tiene un peso de 0 y otros que tiene un peso de 1. Según los resultados, el que es más determinante para predecir si el alumno aprobará o suspenderá la asignatura es el examen *Parcial2* en la mayoría de métodos de ponderación y en algunos métodos *P2* o *P1*.

Finalmente, en la asignatura de TIC, el examen *Parcial1* obtiene mayoritariamente en todos los métodos de ponderación un peso de 0 o casi 0, esto indica que no es determinante para predecir si un alumno aprobará o suspenderá la asignatura. En cambio, *E2* parece ser muy determinante. Este factor también se ha podido observar en la evaluación por rangos.

5. Conclusiones

En este artículo se ha presentado un recurso docente que permite evaluar de manera automática la evolución del alumnado en el aprendizaje de una asignatura y realizar una estimación de la nota final que obtendría al finalizar el curso.

El recurso docente se ofrece al profesorado en forma de una herramienta multiplataforma denominada GRADEFORSEER desarrollada en Java. GRADEFORSEER proporciona una interfaz simple y fácil de usar, que permite predecir en

	InfoGain	Chisquare	Correlation	Relieff	GainRatio	Significance	Simmetrical
DSW	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0
	E3 : 0.0	E3 : 0.0	E3 : 0.49	E3 : 0.08	E3 : 0.0	E3 : 0.0	E3 : 0.0
	P2 : 0.58	P2 : 0.56	P2 : 0.53	P2 : 0.24	P2 : 0.67	P2 : 0.79	P2 : 0.62
	P1 : 0.33	P1 : 0.32	P1 : 0.72	P1 : 0.17	P1 : 0.93	P1 : 0.879	P1 : 0.46
	E2 : 0.0	E2 : 0.0	E2 : 0.23	E2 : 0.0	E2 : 0.0	E2 : 0.0	E2 : 0.0
	P3 : 0.49	P3 : 0.48	P3 : 0.48	P3 : 0.46	P3 : 0.81	P3 : 0.77	P3 : 0.59
	E1 : 0.0	E1 : 0.0	E1 : 0.0	E1 : 0.07	E1 : 0.0	E1 : 0.0	E1 : 0.0
Parcial1: 0.41	Parcial1: 0.49	Parcial1: 0.53	Parcial1: 0.293	Parcial1: 0.94	Parcial1: 0.86	Parcial1: 0.54	
P1	P2 : 0.59	P2 : 0.85	P2 : 0.29	P2 : 0.06	P2 : 0.28	P2 : 0.69	P2 : 0.47
	P3 : 0.61	P3 : 0.53	P3 : 0.53	P3 : 0.26	P3 : 0.33	P3 : 0.44	P3 : 0.50
	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 0.83	Parcial2: 0.80	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0
	P4 : 0.90	P4 : 0.74	P4 : 1.0	P4 : 1.0	P4 : 0.71	P4 : 0.74	P4 : 0.83
	Parcial1: 0.56	Parcial1: 0.98	Parcial1: 0.52	Parcial1: 0.04	Parcial1: 0.56	Parcial1: 0.79	Parcial1: 0.56
P1 : 0.0	P1 : 0.0	P1 : 0.0	P1 : 0.0	P1 : 0.0	P1 : 0.0	P1 : 0.0	
FDT	P2 : 0.33	P2 : 0.12	P2 : 0.25	P2 : 1.0	P2 : 0.0	P2 : 0.0	P2 : 0.17
	Parcial1: 0.85	Parcial1: 0.51	Parcial1: 0.66	Parcial1: 0.0	Parcial1: 0.46	Parcial1: 0.59	Parcial1: 0.67
	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 0.0	Parcial2: 0.198	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0
	P1 : 0.01	P1 : 0.02	P1 : 0.47	P1 : 0.14	P1 : 0.22	P1 : 0.13	P1 : 0.0
P3 : 0.0	P3 : 0.0	P3 : 1.0	P3 : 0.84	P3 : 0.27	P3 : 0.06	P3 : 0.0	
TIC	E1 : 0.0	E1 : 0.0	E1 : 0.0	E1 : 0.0	E1 : 0.0	E1 : 0.0	E1 : 0.0
	E3 : 0.36	E3 : 0.34	E3 : 0.87	E3 : 0.73	E3 : 0.76	E3 : 0.52	E3 : 0.45
	E2 : 1.0	E2 : 1.0	E2 : 1.0	E2 : 1.0	E2 : 1.0	E2 : 1.0	E2 : 1.0
	Parcial1: 0.34	Parcial1: 0.24	Parcial1: 0.54	Parcial1: 0.48	Parcial1: 0.43	Parcial1: 0.19	Parcial1: 0.39
	Parcial2: 0.10	Parcial2: 0.10	Parcial2: 0.20	Parcial2: 0.19	Parcial2: 0.24	Parcial2: 0.18	Parcial2: 0.13

Cuadro 3: Resultados de los pesos sobre cada asignatura por rango.

etapas iniciales-medias del curso la nota final de los alumnos. Disponer de esta herramienta permite, por un lado, reforzar los conocimientos y plantear nuevas actividades a aquellos alumnos que tienen como nota predictiva un suspenso y, por otro lado, motivar a aquellos alumnos que a pesar de tener notas bajas en las primeras actividades del curso obtienen un aprobado según la predicción propuesta por la herramienta.

El núcleo de dicho recurso es un módulo de predicción basado en técnicas de aprendizaje automático que utiliza 87 algoritmos de clasificación procedentes de las librerías CBR-KM y WEKA. Los clasificadores usan como datos de entrenamiento los ficheros con los datos completos de los cursos académicos anteriores de la misma asignatura y como datos a predecir los ficheros con los datos incompletos del curso académico actual.

La evaluación del sistema de predicción se ha realizado con cuatro asignaturas de disciplinas distintas: Diseño de Software (DSW) y Programación I (PI) del grado de informática, Fundamentos de Tecnología (FDT) que se imparte en el grado de Información y Documentación y, finalmente, Tecnologías de la Información y la Comunicación (TIC) impartida en el grado de Lingüística. En general, se han obtenido las mejores predicciones en los resultados evaluados de forma binaria (aprueba o no la asignatura), siendo éstos de un 91,6 %, un 92,2 %, un 91,6 % y un 90,9 % de acierto en DSW, PI, DFT y TIC respectivamente. Además, la evaluación de la capacidad predictiva del sistema de predicción con notas parciales ha mostrado un alto porcentaje de aciertos (un 88,0 % en DSW, un 85,9 % en PI, un 88,7 % en FDT y un 86,8 % en TIC). Nótese que en los resultados sólo se han usado aproximadamente el 60 % de las pruebas del curso académico actual para predecir si un alumno aprueba o no la asignatura. Finalmente, se

han mostrado los resultados de analizar qué actividades evaluativas son las que más influyen en el resultado del alumno y se ha podido contrastar diferentes métodos de ponderación de atributos. Los resultados contrastados denotan que hay actividades más relevantes que otras.

Como trabajo futuro se plantea la integración de GRADE-FORESEER en plataformas de aprendizaje (Moodle), el análisis del uso de políticas alternativas en la generación de la predicción de notas y detectar qué pruebas de evaluación a lo largo del curso inciden más en la nota final.

Agradecimientos

Este trabajo ha sido posible gracias a los proyectos *Módulo pedagógico en un Sistema Tutor Inteligente para predecir la evolución del alumnado* (2014PID-UB030), SGR-623-2014, TIN2012-38603-C02, TIN2011-24220 y TIN2012-38876-C02-02.

Referencias

- [1] P. Fernández Sánchez. *La función tutorial*. Madrid: Castalia-M. E. C, 1991.
- [2] Hendrik Drachsler y Wolfgang Greller. *Turning learning into numbers a learning analytics framework*. Slideshare, 30 de agosto de 2011. Disponible en <http://es.slideshare.net/Drachsler/turning-learning-into-numbers-a-learning-analytics>
- [3] A. Graesser, P. Chipman, B. Haynes y A. Olney. Autotutor: an intelligent tutoring system with mixed-initiative

	InfoGain	Chisquare	Correlation	Relieff	GainRatio	Significance	Simmetrical
DSW	P1 : 0.33	P1 : 0.34	P1 : 0.66	P1 : 0.18	P1 : 0.27	P1 : 0.55	P1 : 0.31
	P2 : 0.39	P2 : 0.54	P2 : 0.57	P2 : 0.11	P2 : 1.0	P2 : 0.97	P2 : 0.63
	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 0.87	Parcial2: 0.95	Parcial2: 1.0
	E1 : 0.0	E1 : 0.0	E1 : 0.0	E1 : 0.240	E1 : 0.0	E1 : 0.0	E1 : 0.0
	E2 : 0.23	E2 : 0.31	E2 : 0.44	E2 : 0.0	E2 : 0.84	E2 : 0.90	E2 : 0.40
	Parcial1: 0.64	Parcial1: 0.90	Parcial1: 0.59	Parcial1: 0.47	Parcial1: 0.88	Parcial1: 1.0	Parcial1: 0.80
	E3 : 0.34	E3 : 0.25	E3 : 0.31	E3 : 0.32	E3 : 0.31	E3 : 0.48	E3 : 0.35
P3 : 0.64	P3 : 0.69	P3 : 0.67	P3 : 0.40	P3 : 0.44	P3 : 0.64	P3 : 0.56	
P1	P3 : 0.47	P3 : 0.55	P3 : 0.56	P3 : 0.07	P3 : 0.17	P3 : 0.23	P3 : 0.36
	P2 : 0.16	P2 : 0.23	P2 : 0.32	P2 : 0.0	P2 : 0.15	P2 : 0.29	P2 : 0.18
	P4 : 0.75	P4 : 0.88	P4 : 1.0				
	P1 : 0.0	P1 : 0.0	P1 : 0.0	P1 : 0.211	P1 : 0.0	P1 : 0.0	P1 : 0.0
	Parcial1: 0.62	Parcial1: 0.64	Parcial1: 0.55	Parcial1: 0.03	Parcial1: 0.28	Parcial1: 0.38	Parcial1: 0.50
	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 0.85	Parcial2: 0.50	Parcial2: 0.67	Parcial2: 0.52	Parcial2: 0.94
FDT	Parcial1: 0.05	Parcial1: 0.63	Parcial1: 1.0	Parcial1: 0.31	Parcial1: 0.59	Parcial1: 0.96	Parcial1: 0.30
	P1 : 0.01	P1 : 0.42	P1 : 0.0	P1 : 0.85	P1 : 0.62	P1 : 0.90	P1 : 0.31
	Parcial2: 0.53	Parcial2: 0.61	Parcial2: 0.59	Parcial2: 0.0	Parcial2: 1.0	Parcial2: 1.0	Parcial2: 1.0
	P3 : 0.0	P3 : 0.0	P3 : 0.56	P3 : 0.95	P3 : 0.74	P3 : 0.76	P3 : 0.41
	P2 : 1.0	P2 : 1.0	P2 : 0.11	P2 : 1.0	P2 : 0.0	P2 : 0.0	P2 : 0.0
TIC	Parcial1: 0.0	Parcial1: 0.0	Parcial1: 0.0	Parcial1: 0.13	Parcial1: 0.0	Parcial1: 0.0	Parcial1: 0.0
	E3 : 0.58	E3 : 0.65	E3 : 0.92	E3 : 0.85	E3 : 0.64	E3 : 0.74	E3 : 0.61
	E1 : 0.12	E1 : 0.11	E1 : 0.09	E1 : 0.0	E1 : 0.27	E1 : 0.26	E1 : 0.17
	Parcial2: 0.25	Parcial2: 0.27	Parcial2: 0.27	Parcial2: 0.22	Parcial2: 0.37	Parcial2: 0.41	Parcial2: 0.30
	E2 : 1.0						

Cuadro 4: Resultados de los pesos sobre cada asignatura por criterio binario.

- dialogue. *IEEE Transactions on Education*, vol. 48, núm. 4, pp. 612–618. Noviembre de 2005.
- [4] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer Peter Reutemann e Ian H. Witten. *The weka data mining software: An update*. SIGKDD Explorations Newsletter, vol. 11, núm. 1, pp. 10–18. Noviembre de 2009.
- [5] Laura Igual Muñoz, Eloi Puertas Prats, Oriol Pujol Vila, Jordi Vitria Marca, Petia Ivanova Radeva, Maria del Pilar Folgueiras Bertomeu, Luis Garrido Ostermann. *Sistema intel·ligent de suport al tutor d'estudis*. Programa Millora Innovació Docent. Universitat de Barcelona, <http://mid.ub.edu/webpamid/content/sistema-intel%20%80%A2ligent-de-suport-al-tutor-d%20%80%99estudis>. 2014.
- [6] Reynol Junco. *The relationship between frequency of facebook use, participation in facebook activities, and student engagement*. *Computers & Education*, vol. 58, núm. 1, pp. 162 – 171. Enero de 2012.
- [7] Mohamed Koutheaïr Khribi, Mohamed Jemni y Olfa Nasraoui. *Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval*. *Educational Technology & Society*, vol. 12, núm. 4, pp. 30–42. 2009.
- [8] J. Marcus. *Here's the New Way Colleges Are Predicting Student Grades*. Time. 10 de Diciembre de 2014 Disponible en <http://time.com/3621228/college-data-tracking-graduation-rates/>.
- [9] Irene Martí Diéguez. *Plataforma de Raonament Basat en Casos per a la gestió del coneixement*. Trabajo Final de Grado, Facultat de Matemàtiques, Universitat de Barcelona. 24 de junio de 2014. Disponible en <http://hdl.handle.net/2445/62504>.
- [10] Claudia Marisa Pagano. *Los tutores en la educación a distancia. un aporte teórico*. *Revista de Universidad y Sociedad del Conocimiento*, vol. 4, núm. 2. 2007. Disponible en <http://www.uoc.edu/rusc/4/2/dt/esp/pagano.pdf>.
- [11] M.M.T. Rodrigo, R.S.J. de Baker, J. Agapito, J. Nabos, M.C. Repalam, S.S. Reyes y M.O.C.Z. San Pedro. The effects of an interactive software agent on student affective dynamics while using an intelligent tutoring system. *IEEE Transactions on Affective Computing*, vol. 3, núm. 2, pp. 224–236. Abril-junio de 2012.
- [12] Maria Salamó Llorente, Inmaculada Rodríguez Santiago, Simone Balocco, Anna Puig Puig, Maite López-Sánchez, Maria Taulé Delor, Ma. Antònia Martí Antonin. *Mòdul pedagògic en un sistema tutor intel·ligent per a predir l'evolució de l'alumnat*. Programa Millora Innovació Docent. Universitat de Barcelona, <http://mid.ub.edu/webpamid/content/m%3%B2dul-pedag%3%B2gic-en-un-sistema-tutor-intel%2%B7ligent-predir-l%20%80%99evoluci%3%B3-de-l%20%80%99alumnat>. 2014.
- [13] Maria Salamó y Maite López-Sánchez. *Adaptive case-based reasoning using retention and forgetting strate-*

gies. Knowledge-Based Systems, vol. 24, núm. 2, pp. 230–247. Marzo de 2011.

- [14] George Siemens. *Learning analytics: The emergence of a discipline*. American Behavioral Scientist, vol. 57 núm. 10, pp. 1380–1400. Octubre de 2013.
- [15] Pei-Chen Sun, Ray J. Tsai, Glenn Finger, Yueh-Yang Chen, Dowming Yeh. *What drives a successful e-learning? an empirical investigation of the critical factors influencing learner satisfaction*. Computers & Education, vol. 50, núm. 4, pp. 1183 – 1202. Mayo de 2008.
- [16] David Vernet, Maria Salamó, Carles Vallespi, Joan Camps, Elisabet Golobardes y Jaume Bacardit. *¿Cómo predecir la evolución del alumno?* En Actas de las VII Jornadas Enseñanzas Universitarias de Informática, Jenui 2001, pp. 329–334. Palma de Mallorca, junio de 2001.



La Dra. *Maria Salamó* es profesora Titular Universitaria e investigadora en la Facultad de Matemáticas de la Universidad de Barcelona (UB), con un sexenio de investigación reconocido. El interés científico de la Dra. Salamó se desarrolla en torno al papel que juega la inteligencia artificial para ayudar a los usuarios a mejorar sus experiencias

en sus tareas diarias. Este interés científico se enfoca actualmente en la línea de investigación de aprendizaje automático y en los sistemas de recomendación dentro de los mundos virtuales. A lo largo de su trayectoria científica, la Dra. Salamó recibió el premio Epon Ibérica en el año 2001 al mejor proyecto de tesis en el ámbito de las Tecnologías de la Información y las Comunicaciones. Además, la Dra. Salamó ha participado a lo largo de su trayectoria científica en 26 proyectos de investigación. En total, la Dra. Salamó tiene más de 60 publicaciones científicas y es miembro del grupo de investigación consolidado Centro de Lenguaje y Computación (2014 SGR 623) y del grupo de innovación docente consolidado INDOMAIN (GIDCUB-13/138).



La Dra. *Inmaculada Rodríguez* La Dra. Inmaculada Rodríguez es profesora Titular de Universidad (TU) e investigadora en la Facultad de Matemáticas de la Universidad de Barcelona (UB), con dos sexenios de investigación reconocidos. Desde 2005 está integrada en el grupo de investigación

WAI de la UB, momento también en el que comenzó una colaboración con investigadores del IIIA del CSIC, en una línea de investigación que combina sistemas multi-agente basados en organizaciones (OCMAS) y mundos virtuales 3D, así como en el diseño de agentes virtuales que asisten a humanos en entornos virtuales 3D. En los últimos años la profesora Rodríguez ha sido responsable en tareas de investigación en los

proyectos nacionales: «Agreement technologies», «Engineering self-virtually embeded systems», «Juegos serios para el entrenamiento de cirugía cardíaca» y « Colaboraciones virtuales robustas». Durante todos estos años ha realizado más de 60 publicaciones en congresos, 13 en revistas, 5 capítulos de libro, además de dirigido 2 tesis doctorales, tesis de máster y numerosos proyectos fin de carrera, muchos de los cuales han estado orientados a temas de investigación en el área de interfaces y juegos serios. La Dra. Rodríguez pertenece a dos grupos de innovación docente de la UB y ha participado en numerosos proyectos de innovación docente que van desde un sistema tutor para la predicción de notas, un programa para el seguimiento y la evaluación del aprendizaje de los trabajos fino de carrera en el grado de ingeniera de Informática, y metodologías de aprendizaje basados en problemas. A lo largo de los últimos años ha formado parte del comité de organización de congresos, ha sido invitada como panelista en congresos, y actualmente forma parte del comité de diferentes congresos, workshops y revistas del área de interacción humano-computador.



La Dra. *Maite López Sánchez* es profesora Titular Universitaria en la Universidad de Barcelona. Asimismo, es la coordinadora en la UB del máster interuniversitario en Inteligencia Artificial (UPC-UB-URV), doctora vinculada al Instituto de Investigación en Inteligencia Artificial (IIIA-CSIC) y miembro de un grupo de innovación docente consolidado INDOMAIN

(GIDCUB-13/138). Anteriormente, la Dra. López fue responsable de investigación en el departamento de innovación de la empresa Intelligent Software Components (iSOCO) e investigadora visitante en el laboratorio de robótica de la Universidad de Southern California (USC). El doctorado lo realizó en 1999 en el área de la Inteligencia Artificial, en el IIIA. A lo largo de su investigación, la Dra. López ha publicado cerca de 100 publicaciones científicas en revistas indexadas, libros y conferencias internacionales de gran prestigio en su área. Cabe también mencionar que algunas de sus contribuciones han sido reconocidas con premios y ha sido investigadora principal y ha participado en proyectos de investigación tanto europeos como nacionales (TIN2012-38876-C02-02, TIN2009-14702-C02-02, CSD2007-0022, MAT-2005-07244-C03-03, IST-1999-19005, IST-2001-34651, IST-2002-507967). Acumula además la dirección de varias tesis doctorales, de entre las que cabe destacar una centrada en la asistencia y la formación de usuarios dentro de los mundos virtuales.



La Dra. *Anna Puig* profesora Titular Universitaria a la Universidad de Barcelona, UB, y tiene reconocidos tres sexenios de investigación. Desde el año 2008 es vicedecana de la Facultat de Matemàtiques de la UB y responsable del grupo de investigación WAI (Volume Visualization and Artificial Intelligence). Ha publicado más

de 60 artículos de investigación en revistas y conferencias internacionales en los ámbitos de Gráficos por Computador e Inteligencia Artificial. Ha participado en diferentes proyectos nacionales de investigación. Forma parte del Grupo de Investigación Consolidado SGR CLIC: Centro de Lenguaje y Comunicación. Su investigación se ha centrado en tres líneas: la visualización 3D y clasificación de imágenes médicas (línea en la que se centró su doctorado), la integración de inteligencia artificial en mundos virtuales 3D y el diseño juegos serios para la educación. Recientemente ha investigado sobre sistemas de recomendación basados en críticas en mundos virtuales 3D en aplicaciones de e-commerce. Actualmente está dirigiendo una tesis relacionada con los juegos y técnicas de gamificación aplicadas al aprendizaje de las matemáticas. Pertenece al grupo de innovación docente INDOMAIN que promueve la difusión de las matemáticas y la informática en secundaria, colaborando en un proyecto de innovación docente para definir tutores inteligentes para mejorar el seguimiento del aprendizaje en asignaturas de grado.



Balocco Simone, desde julio 2015 es profesor Agregado interino de la Universidad de Barcelona, departamento de matemática aplicada y análisis y es investigador en el Centro de Visión por Computador (CVC), Bellaterra. Simone es miembro de un grupo de innovación docente consolidado INDOMAIN (GIDCUB-13/138). Simone se graduó doctor en el 2006 con una tesis co-

supervisada, en Acústica en el laboratorio CREATIS, Universidad Lyon1, Lyon (Francia) y en Electrónica y Telecomunicaciones obtenido al MSD Lab de la Universidad de Florencia (Italia). Desde octubre de 2006 hasta enero 2009, fu contactado como investigador post-doctoral en el laboratorio CISTIB de la Universidad Pompeu Fabra, Barcelona, financiado por el proyecto europeo Aneurist, y al mismo tiempo contratado como profesor asociado de la misma facultad. Sus líneas de in-

vestigación son: el estudio de aprendizaje automático, la inteligencia artificial, de análisis y minería de datos y de reconocimiento de patrones automatizado. Simone ha co-supervisado dos estudiantes de Doctorado: Marina Alberti, que se graduó doctora en el febrero del 2013 y Gerard Blasco, que se graduó doctor en mayo 2015 en la Universitat de Girona, ambos con nota "excelente cum laude". Simone es autor y co-autor de 14 journal internacionales del primer cuartil, en las revistas con más alto índice de impacto de su sector, inventor de 3 patentes internacionales, y co-autor de más que 30 capítulos de libros y conferencias internacionales con proceedings. Ha participado en 14 proyectos nacionales e internacionales (de esos un proyecto integrado europeo) y ha sido IP de un proyecto de acción integrada con la serbia.



Maria Taulé Delor es licenciada en Filología Catalana y doctora en Filología Románica, especialidad de Lingüística Computacional, por la Universidad de Barcelona. Es profesora contratada del departamento de Lingüística General de la misma Universidad des de 1999. Además, es miembro del grupo de investigación CLiC (Centre de Llenguatge i Computació, 2014 SGR

623) de la Universidad de Barcelona. Ha participado en diversos proyectos, nacionales –WordNet del Español y WordNet Catalán, CESS-ECE, Lang2World, CIInt, TexMess 2.0, AnCora-Net, Araknion, Iarg-AnCora, DIANA, etc.– e internacionales –Acquilex, EuroWordNet, Senseval, SemEval–, en los cuales ha desarrollado su actividad investigadora, relacionada con las tecnologías del lenguaje humano y, especialmente, con la semántica léxica, la lingüística de corpus y el desarrollo de recursos de ingeniería lingüística, básicamente para las lenguas castellana, catalana e inglesa. Para más información consulte su página web: <http://clic.ub.edu/en/users/mariona-tauló-delor>.



© 2016 M. Salamó, I. Rodríguez, M. López-Sánchez, A. Puig, S. Balocco, M. Taulé. Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional que permite copiar, distribuir y comunicar públicamente la obra en cualquier medio, sólido o electrónico, siempre que se acrediten a los autores y fuentes originales y no se haga un uso comercial.