

# Detección precoz de dificultades en el aprendizaje. Herramienta para la predicción del rendimiento de los estudiantes

## Early detection of learning difficulties. Tool for predicting student performance

Carlos J. Villagr -Arnedo, Francisco J. Gallego-Dur n, Fara n Llorens-Largo, Patricia Compa n-Rosique, Rosana Satorre-Cuerda, Rafael Molina-Carmona  
{villagra, fgallego, faraon, patricia, rosana, rmolina}@dccia.ua.es

Departamento de Ciencia de la Computaci n e Inteligencia Artificial  
Universidad de Alicante  
Alicante, Espa a

**Resumen-** Inspirados por las estrategias de detecci n precoz aplicadas en medicina, proponemos el dise o y construcci n de un sistema de predicci n que permita detectar los problemas de aprendizaje de los estudiantes de forma temprana. Partimos de un sistema gamificado para el aprendizaje de L gica Computacional, del que se recolectan masivamente datos de uso y, sobre todo, resultados de aprendizaje de los estudiantes en la resoluci n de problemas. Todos estos datos se analizan utilizando t cnicas de Machine Learning que ofrecen, como resultado, una predicci n del rendimiento de cada alumno. La informaci n se presenta semanalmente en forma de un gr fico de progresi n, de f cil interpretaci n pero con informaci n muy valiosa. El sistema resultante tiene un alto grado de automatizaci n, es progresivo, ofrece resultados desde el principio del curso con predicciones cada vez m s precisas, utiliza resultados de aprendizaje y no solo datos de uso, permite evaluar y hacer predicciones sobre las competencias y habilidades adquiridas y contribuye a una evaluaci n realmente formativa. En definitiva, permite a los profesores guiar a los estudiantes en una mejora de su rendimiento desde etapas muy tempranas, pudiendo reconducir a tiempo los posibles fracasos y motivando a los estudiantes.

**Palabras clave:** *detecci n precoz, predicci n, evaluaci n formativa, motivaci n.*

**Abstract-** Inspired by the early detection strategies applied in medicine, we propose the design and construction of a prediction system to early detect learning problems of students. A gamified system to learn Computational Logic is the starting point from which a massive set of usage data and learning outcomes about problem solving is collected. All these data are analysed using Machine Learning techniques. As a result, a prediction of the performance of each student is obtained. The information is weekly presented as a progression chart, which is easily interpretable and contains valuable information. The resulting system has a high degree of automation, is progressive, provides results from the beginning of course with increasingly accurate predictions, uses learning outcomes as well as usage data, allows the evaluation and prediction of the acquired skills and abilities, and contributes to a truly formative assessment. In short, it allows teachers to guide students in their performance improvement from very early stages and can redirect possible failures in time and motivate students.

**Keywords:** *early detection, prediction, formative assessment, motivation.*

### 1. INTRODUCCI N

La detecci n precoz es un t rmino m dico empleado para referirse a las estrategias utilizadas en una poblaci n para identificar una enfermedad que, en sus primeros estadios, es asintom tica. El objetivo de estas acciones es, en  ltima instancia y gracias a la pronta intervenci n, reducir la mortalidad y el padecimiento de una enfermedad. Estos conceptos propios de la epidemiolog a nos han servido de inspiraci n para este trabajo. Aunque los problemas de aprendizaje de los estudiantes no constituyen, desde luego, una enfermedad, podemos establecer un cierto paralelismo entre la labor de un m dico, centrada en detectar y curar las enfermedades, y la de un docente, que persigue identificar las carencias en el aprendizaje de sus estudiantes y ofrecerles una soluci n.

Mejorar el rendimiento de los estudiantes, conocer el progreso real de estos y tratar de predecir los resultados en las etapas m s tempranas del proceso docente puede ser de una extraordinaria importancia para actuar de forma precoz y atajar los problemas de ra z. Esto ha motivado varias l neas de investigaci n en diferentes campos que tienen como objetivo encontrar la forma de predecir los resultados de aprendizaje. Centr ndonos en el  mbito de las herramientas computacionales, en la literatura existen algunos sistemas de predicci n, entre ellos los que utilizan diversas t cnicas de inteligencia artificial y miner a de datos (Canaleta, Sol  & Navarro, 2012; Romero, Zafra, Gibaja, Luque & Ventura, 2012; Vernet, Salam , Vallesp , Camps, Golobardes & Bacardit, 2001), y otros m todos que son combinados con la gamificaci n (Illanas, Calvo, Gallego & Llorens, 2013; Illanas, Llorens, Molina, Gallego, Compa n, Satorre & Villagr , 2014). El problema es que la mayor a tiende a usar modelos basados  nicamente en encuestas, informaci n y calificaciones previas al curso y algunas notas puntuales, y los utilizan bien para detectar las metodolog as docentes que proporcionan los mejores resultados, bien para predecir las notas finales, pero s lo a nivel de aprobado o suspenso.

Un ejemplo interesante es el sistema descrito por Vernet *et al.* (2001), un m todo de predicci n combinado con un sistema de tutorizaci n inteligente. A pesar de que sus resultados son buenos, el m todo dise ado no estaba pensado

para ser mantenido en el tiempo, ni para hacer predicciones progresivas basadas en información incremental.

En los últimos años se han producido avances importantes en el campo de Learning Analytics y, particularmente, en el uso de las técnicas de Machine Learning en la predicción de los resultados de aprendizaje. Kotsiantis (2012) hace una interesante revisión de la que se extrae la conclusión de que la mayoría de los modelos se construyen en un momento dado a partir de los datos disponibles y se utilizan para predecir el rendimiento futuro de los estudiante en forma de notas o calificaciones. Por ejemplo, Schalk, Wick, Turner y Ramsdell (2011) presentan un sistema predictivo para estudiantes de cursos introductorios de matemáticas y física. El sistema utiliza la técnica de Random Forests para analizar datos procedentes de cursos anteriores del examen Standard Admission Test y, a partir de ellos, estiman qué estudiantes se encuentran en situación de riesgo. Otro ejemplo interesante se encuentra en (Huang & Fang, 2013), donde los autores comparan cuatro modelos de Machine Learning para predecir el desempeño del estudiante. Los mejores resultados predictores para estudiantes individuales se obtienen utilizando Support Vector Machines. La investigación se centra en la calidad de las predicciones para un instante concreto en el tiempo. Por otro lado, Lykourantzou, Giannoukos, Mpardis, Nikolopoulos y Loumos (2009) presentan un modelo que logra una predicción precisa en una etapa temprana de un curso de eLearning utilizando redes neuronales feed-forward. Sin embargo, el método propuesto no logra predecir con precisión el rendimiento de ciertos alumnos específicos. En general, todos estos métodos, aun obteniendo resultados satisfactorios en muchos casos, no tienen como horizonte mantener el sistema en el tiempo ni realizar predicciones con datos incrementales.

Un caso distinto es el del sistema de alerta temprana propuesto por Hu, Lo & Shih (2014), pensado para ayudar a identificar a los estudiantes en riesgo y predecir su aprendizaje mediante el análisis de rendimiento del portafolio docente registrado en un sistema de gestión de aprendizaje (LMS, por sus siglas en inglés). A diferencia de otras propuestas, el sistema es automático e incremental, ofreciendo información en tres ocasiones durante el curso. Su principal limitación se debe a que utiliza como entrada exclusivamente los datos de uso del LMS, por lo que sus conclusiones se centran más en la actividad de los estudiantes que en el aprendizaje efectivo y las competencias adquiridas.

Inspirados por algunos de estos trabajos y motivados por intentar resolver algunas de las carencias que presentan, nos proponemos como objetivo de este trabajo diseñar y construir un sistema predictivo, de ayuda para el docente y para los propios estudiantes, que permita detectar de forma precoz los problemas de aprendizaje, que presente un alto grado de automatización, que ofrezca predicciones sobre el aprendizaje efectivo del estudiante y las competencias adquiridas, que sea perdurable en el tiempo y que utilice, en cada etapa, datos nuevos acumulados a la experiencia adquirida. En definitiva, un sistema que evolucione para obtener cada vez un mejor diagnóstico del proceso de aprendizaje de nuestros estudiantes. Nuestra propuesta es una herramienta automática, que toma como base PLMan (Castel, Gallego, Pomares, Suau, Villagrà & Cortés, 2009), un sistema gamificado personalizado para el aprendizaje de Lógica Computacional, creado hace unos años y probado exhaustivamente. A partir de su uso por los

estudiantes, realiza una recolección masiva de datos y un análisis de los mismos utilizando técnicas de Machine Learning. Como resultado, el sistema ofrece para cada alumno una predicción de su rendimiento, presentada de forma semanal y acumulada en forma de un gráfico de progresión. El resultado es de una ayuda inestimable para descubrir patrones de aprendizaje y actuar sobre ellos.

En el apartado 2 presentamos el contexto de esta propuesta, presentando el entorno docente en el que se engloba y sus objetivos. En el apartado 3 describimos con detalle las características de nuestra herramienta y cómo se ha llevado a cabo la experiencia. Los resultados del trabajo se presentan en el apartado 4. Las conclusiones de la experiencia conforman el último apartado.

## 2. CONTEXTO

La experiencia que se presenta se enmarca en la docencia de una asignatura de primer curso de los Grados en Ingeniería Informática e Ingeniería Multimedia, cuyo objetivo es introducir a los estudiantes en la materia de Lógica Computacional. Aunque la experiencia está muy centrada en las características de esta asignatura, su aplicación a otras materias es posible. En la sección de conclusiones daremos algunas pautas sobre cómo extender el modelo a otros ámbitos.

La finalidad de las prácticas de Lógica Computacional es conseguir que los estudiantes sean capaces de resolver problemas relacionados con la materia utilizando el lenguaje de programación Prolog. Para ello, hemos diseñado y construido un sistema de aprendizaje personalizado y gamificado denominado PLMan (Castel et al., 2009), como ya se mencionó en la introducción. La gamificación está teniendo últimamente un papel destacado en la educación por incluir varios aspectos claves en la mejora de la motivación: retroalimentación inmediata, dinámica de juegos, aprendizaje personalizado y evaluación automática (Calderón & Ruiz, 2014; Espinilla, Palomares, Martínez, Pagola & Bustince, 2010; Gallego, Villagrà, Satorre, Compañ, Molina & Llorens, 2014).

PLMan tiene el aspecto de un videojuego sencillo, en el que los estudiantes programan la inteligencia artificial de un personaje similar a Pac-Man, usando una base de conocimientos de Prolog. Cada ejercicio se presenta en forma de un laberinto o mapa, con trampas y enemigos, que los estudiantes deben resolver mediante la programación, a través de un conjunto de reglas simples en Prolog, del comportamiento inteligente del personaje, cuyo objetivo es comer el máximo número de cocos evitando a los enemigos. La complejidad de los laberintos aumenta y en cada nivel requiere de los estudiantes mejores habilidades en la programación y mayor creatividad.

Los mapas, de los que ya se han creado más de 400 variantes organizadas en 4 categorías correspondientes a sendas fases y 5 niveles de dificultad, tienen diseños distintos, diferentes objetos que se pueden coger y usar, varios enemigos y obstáculos a superar e incluso problemas que solucionar. Los estudiantes deben superar las 4 fases y una fase final denominada examinador, para conseguir la máxima calificación en el sistema. En cada fase, tienen que resolver de 1 a 5 laberintos diferentes, dependiendo de la fase. Primero, seleccionan el nivel de dificultad preferido (de 1 a 5) y el sistema les asigna un mapa aleatorio, que es diferente para cada estudiante del mismo grupo de prácticas. Entonces, deben programar offline

la inteligencia artificial que permite superarlo y subir la solución al sistema. No existe un límite establecido de entregas para un determinado laberinto ni se penaliza el error. Simplemente, si no obtienen una solución satisfactoria, pueden continuar desarrollando y probando su solución hasta que logran superarlo. También pueden seguir su propio ritmo eligiendo los niveles de dificultad con los que se sientan más cómodos, teniendo en cuenta que cuanto mayor sea el nivel de dificultad, mayor será la nota que acumularán a la calificación final. Cuando logran más del 75% de cocos comidos en un laberinto, consiguen desbloquear el siguiente y continúan seleccionando dificultad. Los estudiantes también pueden detenerse cuando lo consideren oportuno: por ejemplo, si llegan a la tercera fase y tienen una nota acumulada de un 6,5, pueden decidir no continuar resolviendo mapas.

El sistema gamificado PLMan persigue, como primer objetivo, mejorar la motivación de los estudiantes. Este aspecto constituye uno de los principales retos del proceso docente, como demuestra el interés que la mejora de la motivación, particularmente la intrínseca, despierta entre los investigadores (Alonso, 1995; Rinaudo, Chiecher & Donolo, 2003).

La evaluación es otro de los objetivos fundamentales de PLMan. Frente a los modelos docentes clásicos, que proponen un ritmo homogéneo de aprendizaje, cuya forma de evaluación es puramente informativa y únicamente clasifica a los estudiantes según sus resultados, nosotros proponemos un modelo instruccional, centrado en el estudiante, con una evaluación heterogénea, formativa, personalizada y con retroalimentación inmediata (Merrill, 2009; Reigeluth, 2012). La enorme carga de trabajo que puede suponer una evaluación de estas características para el profesor, se ve mitigada gracias al uso de las nuevas tecnologías. Una de las claves del éxito de PLMan está en la automatización, que nos permite atender las necesidades de muchos estudiantes en mejores condiciones que si hemos de hacerlo con métodos tradicionales.

El tercer objetivo de la experiencia es proporcionar a los estudiantes y a los profesores una herramienta que les permita conocer durante el desarrollo de proceso docente la situación del aprendizaje efectivo de los estudiantes y, sobre todo, estimar su evolución futura, con la finalidad de realizar una detección precoz eficaz de los posibles problemas. Buscamos una herramienta gráfica y sencilla pero eficaz, que permita descubrir patrones de aprendizaje y actuar sobre ellos. Esta herramienta centra el desarrollo de los siguientes apartados y constituye la principal aportación de este trabajo.

### 3. DESCRIPCIÓN

Un sistema como PLMan genera una enorme cantidad de datos relacionados tanto con la interacción y uso del sistema como con los resultados del proceso de aprendizaje. No podemos desaprovechar este ingente flujo de datos que, convenientemente tratado, puede transformarse en valiosísima información, útil tanto para el profesor como para el estudiante.

#### A. Recolección de datos

La primera cuestión a tratar es el tipo de datos que es útil analizar. Consideramos una serie de eventos concretos que ocurren durante la interacción de los usuarios con el sistema. Todos estos sucesos se registran en una base de datos de eventos, incluyendo la fecha y hora en que se produjeron y su información relacionada dependiendo del tipo de evento: accesos al sistema, visualización de resultados, selección de

dificultades, descargas de laberintos, entrega de soluciones correctas e incorrectas, descarga de logs de ejecución, etc. Este grupo simple de eventos proporciona exclusivamente datos de uso del sistema.

Junto con estos eventos, se recopilan datos sobre el estado en el que se encuentra el estudiante: mapas resueltos, notas obtenidas y fases superadas. Estos datos proporcionan información sobre el desarrollo del proceso de aprendizaje y las competencias adquiridas. Esta información sobre el aprendizaje efectivo es muy valiosa y supone una aportación importante frente a otros sistemas de Learning Analytics que únicamente utilizan datos de uso.

Finalmente, a partir de todos estos datos sobre el uso del sistema y sobre el aprendizaje del estudiante, proponemos el siguiente conjunto de características diseñadas para su uso en la herramienta predictiva:

- Número de visitas al frontal
- Número de descargas de mapas
- Número de entregas por fase
- Promedio de notas por fase (porcentaje)
- Tiempo empleado en superar cada fase (segundos)
- Tiempo empleado en superar cada mapa (segundos)
- Niveles de dificultad seleccionados por mapa

#### B. Algoritmo de predicción

Las características extraídas del conjunto de datos, representan la entrada para un algoritmo de Machine Learning que predice el rendimiento que van a obtener los estudiantes al final del cuatrimestre. Este algoritmo está basado en el método Support Vector Machines (SVM) con acoplamiento por parejas (Cortes & Vapnik, 1995). Se trata de una técnica que funciona muy bien con conjuntos de datos como el que se analiza en este trabajo. El sistema de predicción clasifica el rendimiento esperado de los estudiantes en una de tres posibles clases, dependiendo de la nota final prevista sobre 10 puntos:

- Estudiantes de rendimiento alto (nota final esperada  $> 8,05$ )
- Estudiantes de rendimiento medio (nota final esperada entre 5,75 y 8,05)
- Estudiantes de rendimiento bajo (nota final esperada  $< 5,75$ )

El principal motivo de esta división en tres clases es la optimización del algoritmo de clasificación. Con los datos disponibles (se han analizado 336 estudiantes en esta primera experiencia), considerar más de 3 clases supondría tener en cuenta muy pocos ejemplos de cada clase, lo que minaría las posibilidades del algoritmo.

Después de seleccionar el mejor conjunto de parámetros para la SVM, se entrenan 336 SVMs, una para cada estudiante. Para cada SVM se usan 335 ejemplos de entrenamiento, dejando fuera al estudiante que va a ser analizado. Entonces, se obtienen las predicciones semanales para cada estudiante, usando sus correspondientes SVMs. Estas predicciones se añaden a la base de datos de predicciones, donde se almacenan las de las semanas anteriores. Por último, estas predicciones se proporcionan a estudiantes y profesores en forma de progresión estimada del rendimiento de los estudiantes en el tiempo.

El sistema gamificado se articula a través de un portal web que da acceso al sistema PLMan y donde los estudiantes

realizan sus entregas y van generando los eventos y los datos de progresión. El sistema registra dichos eventos y los procesa obteniendo las características con las que opera el sistema predictivo.

#### 4. RESULTADOS

El sistema propuesto ha sido implementado en el primer cuatrimestre del curso actual, empezando su actividad en octubre de 2014 y finalizando en enero de 2015. Había alrededor de 400 estudiantes registrados, de los que 336 participaron activamente en las clases de prácticas usando el sistema de forma habitual.

Durante la asignatura se imparten 13 clases de prácticas efectivas, de las que 11 son útiles a nivel de seguimiento (en las dos primeras se imparte una introducción general a Prolog y PLMan). De esta forma, el sistema proporciona a los profesores 10 semanas de predicciones, y termina en la semana 11 proporcionando la nota final a los estudiantes.

La salida del sistema de predicción es, para cada estudiante y cada una de las 10 semanas, 3 valores de probabilidad estimados, que representan, respectivamente, la probabilidad estimada de que el estudiante tenga, al final del curso, un rendimiento alto, un rendimiento medio o un rendimiento bajo. La mayor probabilidad de las 3, marcará la clase en la cual el estudiante se clasifica. Al acabar el curso y conocer finalmente la calificación del estudiante, podemos calcular la tasa de aciertos del sistema predictivo para cada semana. En la figura 1 se presenta una gráfica con la tasa media de aciertos cada semana, considerando todos los estudiantes. Esta figura corrobora lo que ya podíamos intuir: cuantos más datos se analizan, mejor es la predicción. Los resultados en las semanas 8 a 10 resultan, no obstante, llamativos, pues existe una tasa de aciertos relativamente baja (en torno al 70%) cuando los estudiantes han realizado su trabajo casi completamente y debería ser sencillo predecir la clasificación correcta. Esto se debe a la separación en clases que se ha realizado: las clases son discretas, mientras que la calificación obtenida está en un espacio continuo. De esta forma, la consistencia de la clasificación para las calificaciones que se encuentran en la frontera de dos clases, es mucho menor. Esto se corrobora si tenemos en cuenta la segunda clase por probabilidad: en la inmensa mayoría de los casos en los que la primera clase es errónea, la segunda opción es la correcta y con un margen de error muy estrecho (normalmente  $<0.05$ ). Un análisis detallado de estos resultados probabilísticos va más allá del alcance de este trabajo. Además, esta circunstancia no afecta a nuestro principal objetivo, como veremos a continuación.

La idea principal detrás de este trabajo es poder dar la información de tendencias a través de predicciones progresivas. Para conseguir este objetivo, el sistema produce gráficas probabilísticas evolutivas como las de las figuras 2, 3 y 4. Estas gráficas muestran las probabilidades estimadas de que un estudiante tenga un rendimiento alto (verde/arriba), medio (azul/centro) o bajo (rojo/abajo). Como se mencionó antes, estas probabilidades se calculan cada semana del curso y se acumulan en el gráfico, mostrando las tendencias de los estudiantes después de las 10 semanas efectivas del curso. Lógicamente, cada semana se muestra el resultado obtenido con los datos recogidos hasta esa semana.

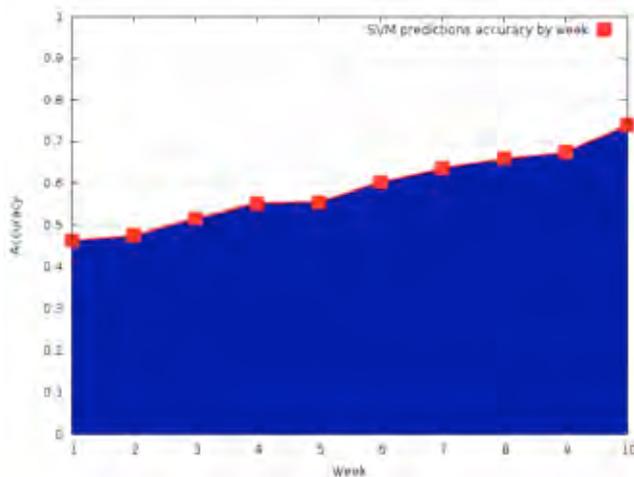


Figura 1: Tasa media de predicciones acertadas cada semana

Los tres estudiantes representados en las figuras 2, 3 y 4 son tres ejemplos típicos de cada categoría: El primer caso (figura 2) es el de un estudiante de rendimiento alto (obtuvo de nota final 9,75 sobre 10), el segundo (figura 3) es un estudiante de rendimiento bajo (consiguió una nota final de 1) y el tercero (figura 4) es un estudiante de rendimiento medio (logró una calificación final de 6,4). Analizando las gráficas de las figuras 2 y 3, parece claro que hay dos tendencias ascendentes y descendentes, en las barras de la parte superior (verde) y en las de la inferior (rojo). Estos estudiantes muestran muy claramente su tendencia entre las semanas 4 y 6. Además, el estudiante de la figura 3 sería un buen ejemplo de detección temprana de un problema, que debería motivar un análisis más detallado por parte del profesor para encontrar una solución. En contraste a esto, entender el patrón que sigue el estudiante de la figura 4 es algo más complejo. Hasta la semana 5, el rendimiento del estudiante tiende a ser medio, pero, de repente, cambia a alto para volver de nuevo a medio en la semana 9. Esto se debe probablemente a que el estudiante resuelve muchos mapas entre las semanas 5 y 6, y entonces se conforma con la nota obtenida y ya no hace nada más. Por ello, en la semana 6 el sistema cree que el estudiante va a hacer muchas cosas y llegará a ser de rendimiento alto, pero esta probabilidad decrece a medida que el esfuerzo vuelve a ser solamente ocasional y no continuo.

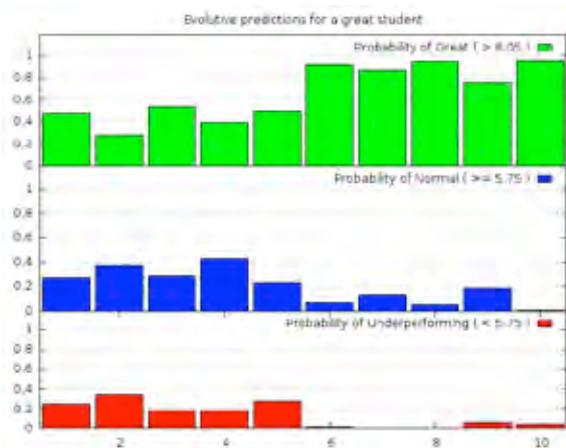


Figura 2: Predicción para un estudiante de rendimiento alto

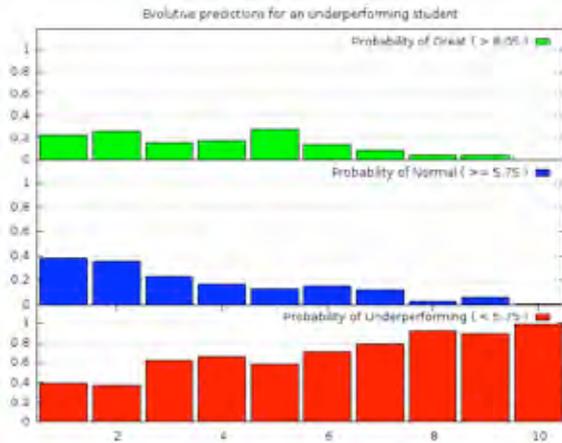


Figura 3: Predicción para un estudiante de rendimiento bajo

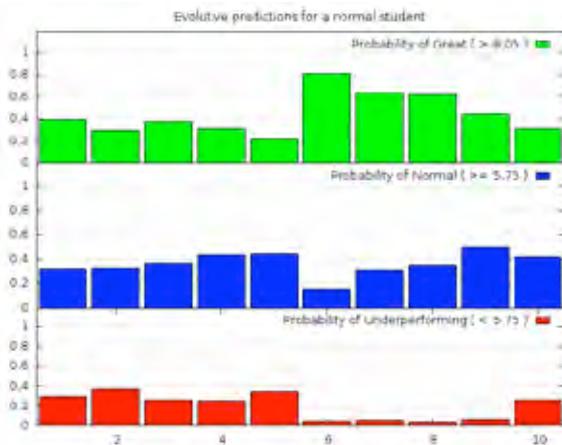


Figura 4: Predicción para un estudiante de rendimiento medio

Este simple ejemplo de análisis muestra la capacidad que las gráficas de predicción progresiva pueden aportar al profesor, permitiéndole detectar patrones de tendencia. Esta innovación les hace capaces de actuar sobre los estudiantes en etapas tempranas y medias de su proceso de aprendizaje, mejorando la posibilidad de guiar a los estudiantes hacia el éxito.

## 5. CONCLUSIONES

Los tres objetivos principales que nos habíamos marcado en este trabajo son mejorar la motivación de los estudiantes, conseguir una verdadera evaluación formativa y obtener una herramienta para conocer y predecir el progreso en el aprendizaje de los estudiantes. Empezando por el último objetivo, hemos construido un sistema automático de recogida y de análisis de datos que nos proporciona información muy valiosa sobre la evolución del alumno y nos permite detectar problemas en el aprendizaje. Las principales aportaciones son:

- El sistema está completamente automatizado, siendo la recogida de datos y extracción de características totalmente transparente para estudiantes y profesores. De esta forma no interfiere en el proceso de aprendizaje de los estudiantes ni incrementa la carga de trabajo para el profesor.
- Se trata de un sistema progresivo, en el que se ofrecen resultados semana a semana, con lo que es posible una reacción a tiempo a los problemas de aprendizaje. La

recolección incremental de datos hace que las predicciones sean cada vez más precisas (para un mismo estudiante a lo largo del curso pero también para el sistema global curso tras curso).

- Es un sistema fácilmente sostenible en el tiempo: nos planteamos utilizarlo cada curso, con poco esfuerzo para su adaptación (lógicamente se irán introduciendo mejoras) y, sobre todo, prevemos conseguir mejores predicciones conforme vayamos aumentando la base de datos disponible.
- A diferencia de otros sistemas de Learning Analytics, no utiliza sólo datos de uso del sistema sino que utiliza datos de resultados de aprendizaje, lo que permite evaluar y hacer predicciones sobre las competencias y habilidades adquiridas.
- Los resultados se presentan en forma de sencillas gráficas, fáciles de interpretar pero con información muy valiosa.

Como resultado de estas características, contribuimos a una evaluación de mayor calidad, que es el segundo de los grandes objetivos de este trabajo. La automatización de tareas repetitivas nos permite dedicar más tiempo a evaluar otros aspectos, resultando en una evaluación más inmediata, más eficaz, más justa y que tenga en cuenta capacidades y conocimientos que antes no se podían valorar. En definitiva, caminamos hacia una evaluación realmente formativa.

Este sistema capacita a los profesores para guiar a los estudiantes en una mejora de su rendimiento desde etapas muy tempranas en su evolución en el sistema, pudiendo reconducir a tiempo los posibles fracasos y motivar a los estudiantes con perspectivas mejores. Esto nos lleva al primero de los objetivos: la mejora de la motivación, no sólo de forma extrínseca (por la obtención de la recompensa en forma de mejor calificación), sino también de forma intrínseca, a través de la iniciativa propia del estudiante.

La extensión de este trabajo a otros ámbitos es posible, aunque puede requerir un esfuerzo importante. La plataforma que da soporte al sistema (que incluye la gestión de usuarios, la entrega de soluciones, la recogida de datos...) es fácilmente configurable. El sistema predictivo basado en técnicas de Machine Learning, también pensamos que puede ser adaptado con poca dificultad a otros contextos, ajustándolo a los datos disponibles. Sin embargo, la adaptación del sistema gamificado que da soporte a la herramienta de predicción no es tan inmediata. Pensamos que, en este caso, la clave está en diseñar un sistema automático de resolución de problemas relacionados con la asignatura en cuestión, que proporcione los datos a la herramienta de predicción. Aunque el diseño de este sistema dependerá totalmente de la materia objeto de aprendizaje, deberá tener ciertas características mínimas: clasificación de los problemas en niveles de dificultad, uso progresivo del sistema con incremento de dificultad y corrección automática de las soluciones. De esta manera será posible obtener datos sobre los resultados de aprendizaje, que son los que confieren el carácter innovador a nuestra propuesta.

En los próximos cursos planeamos mejorar el sistema de predicción en tres aspectos fundamentales: refinando las características de entrada del algoritmo de predicción e investigando nuevos atributos que permitan mejorar la predicción; incorporando nuevas gráficas que describan mejor la situación del proceso de aprendizaje y permitan detectar

otros problemas; y, sobre todo, y volviendo al símil médico con el que empezamos, tratando no sólo de diagnosticar el problema sino también de obtener del sistema una aproximación a cuál es el tratamiento adecuado para resolverlo.

#### REFERENCIAS

- Alonso, J. (1995). *Motivación y aprendizaje en el aula. Cómo enseñar a pensar*. Madrid, España: Santillana.
- Calderón, A. y Ruiz, M. (2014). Evaluación automática en dirección y gestión de proyectos software a través de un juego basado en simulación. En *Actas del Simposio-Taller de las XX Jornadas de la Enseñanza Universitaria de la Informática Jenúí*, Oviedo, España.
- Canaleta, X., Solé, X. y Navarro, J. (2012). Herramienta de soporte a la evaluación del aprendizaje y gestión docente. En *Actas de las XVIII Jornadas de la Enseñanza Universitaria de la Informática Jenúí*, Ciudad Real, España.
- Castel, M.J., Gallego, F., Pomares, C., Suau, P., Villagrà, C. y Cortés, S. (2009). e-Valoración en tiempo real. En *Actas de las XV Jornadas de la Enseñanza Universitaria de la Informática Jenúí*, Barcelona, España.
- Cortes, C., Vapnik, V. (1995). Support-vector networks. *Machine Learning* 20(3), 273–297.
- Espinilla, M., Palomares, I., Martínez, L., Pagola, M. y Bustince, H. (2010). Desarrollo y empleo de juegos educativos on-line destinados al auto-entrenamiento y auto-evaluación. En *Actas de las XVI Jornadas de la Enseñanza Universitaria de la Informática Jenúí*, Santiago de Compostela, España.
- Gallego, F., Villagrà, C., Satorre, R., Compañ, P., Molina, R., y Llorens, F. (2014). Panoràmica: serious games, gamification y mucho más. *ReVisión (Revista de Investigación en Docencia Universitaria de la Informática)* 7(3), 13-23.
- Hu, Y.H., Lo, C.L. y Shih, S.P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior* 36, 469-478.
- Huang, S. y Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education* 61(0), 133-145.
- Illanas, A., Calvo, J.R., Gallego, F. y Llorens, F. (2013). Predicting Student performance in translating foreign languages with a serious game. En *Proceedings 7th International Technology, Education and Development Conference, INTED*. Valencia, España.
- Illanas, A., Llorens, F., Molina, R., Gallego, F., Compañ, P., Satorre, R. y Villagrà, C. (2014). ¿Puede un videojuego ayudarnos a predecir los resultados de aprendizaje? En *Actas I Congreso de la Sociedad Española para las Ciencias del Videojuego CoSECiVi*. Barcelona, España.
- Kotsiantis, S.B. (2012). Use of machine learning techniques for educational proposes: A decision support system for forecasting students' grades. *Artificial Intelligence Review* 37(4), 331-344.
- Lykourentzou, I., Giannoukos, I., Mparadis, G., Nikolopoulos, V., Loumos, V. (2009). Early and dynamic student achievement prediction in e-learning courses using neural networks. *J. Am. Soc. Inf. Sci. Technol.* 60(2), 372-380.
- Merrill, M.D. (2009). First principles of instruction. En C. M. Reigeluth & A. A. Carr-Chellman (Eds.), *Instructional-design theories and models: Building a common knowledge base* 3, 41-56.
- Reigeluth, C. (2012). Teoría instruccional y tecnología para el nuevo paradigma de la educación. En *RED, Revista de Educación a Distancia* 32, recuperado en <http://www.um.es/ead/red/32>
- Rinaudo, M.C., Chiecher, A. y Donolo, D. (2003). Motivación y uso de estrategias en estudiantes universitarios. Su evaluación a partir del Motivated Strategies Learning Questionnaire. En *Anales de psicología* 19(1), 107-119.
- Romero, C., Zafra, A., Gibaja, E., Luque, M. y Ventura, S. (2012). Predicción del rendimiento académico en las nuevas titulaciones de grado de la EPS de la Universidad de Córdoba. En *Actas del Simposio-Taller de las XVIII Jornadas de la Enseñanza Universitaria de la Informática Jenúí*. Ciudad Real, España.
- Schalk, P., Wick, D., Turner, P., Ramsdell, M. (2011). Predictive assessment of student performance for early strategic guidance. In *Frontiers in Education Conference (FIE)*, Rapid City, Estados Unidos de América.
- Vernet, D., Salamó, M., Vallespí, C., Camps, J., Golobardes, E. y Bacardit, J. (2001). ¿Cómo predecir la evolución de un alumno? En *Actas de las VII Jornadas de la Enseñanza Universitaria de la Informática Jenúí*, Palma de Mallorca, España