

Simposio

TÍTULO DEL SIMPOSIO: *Aprendizaje personalizado en Educación Superior: Aplicación de técnicas de inteligencia artificial y de minería de datos*

Coordinadora: Dra. María Consuelo Sáiz Manzanares

e-mail: mcsmanzanares@ubu.es

Institución: Universidad de Burgos

RESUMEN GENERAL DEL SIMPOSIO

Resumen:

Estudios recientes en Educación Superior indican que las metodologías docentes tienen que dirigirse a formas más personalizadas de interacción entre el estudiante y el docente. Los entornos de aprendizaje en la sociedad del siglo XXI han cambiado. En la actualidad un porcentaje alto de la docencia, ya sea presencial o virtual, se efectúa a través de Sistemas de Gestión del Aprendizaje también denominados Learning Management Systems (LMS), como por ejemplo Moodle. Los LMS facilitan la personalización y el respeto al ritmo de aprendizaje de los estudiantes, lo que potencia el desarrollo de la auto-regulación del aprendizaje (SRL) y el incremento de la motivación. Todo ello, va a evitar el abandono en las titulaciones y a aumentar la consecución de aprendizajes eficaces. Recientes estudios indican que el seguimiento personalizado por parte del docente del proceso de aprendizaje de los estudiantes predice los resultados de aprendizaje en un 61.3% y los patrones de conductuales de aprendizaje de estos en los LMS en un 56.1%. Asimismo, la utilización de los LMS incrementa la utilización de estrategias metacognitivas, aspecto relevante ya que éstas explican el 40% de los resultados de aprendizaje. Si bien, la información que se registra en los LMS puede llegar a ser abrumadora. Es por lo que, se hace necesario emplear para su análisis técnicas de minería de datos (Educational Data Mining-EDM-). Las EDM pueden utilizarse en categorías de aprendizaje automático supervisado (clasificación o predicción) y/o no supervisado (agrupamiento o clustering). En este Simposio se presentan los resultados de la aplicación de EDM en cinco estudios con estudiantes universitarios. El primero de ellos analiza la predicción de la variable “trabajo colaborativo en LMS” sobre los resultados de aprendizaje y los patrones conductuales de aprendizaje en la plataforma en dos modalidades de Blended-Learning (B-Learning) Replacement Blend (RB) vs. Supplemental Blend (SB). El segundo, estudia el valor predictivo de los clústeres de agrupamiento respecto del rendimiento académico y de los patrones efectivos de aprendizaje. El tercer trabajo, estudia si los conocimientos previos de los estudiantes y las estrategias metacognitivas son predictores de los resultados de aprendizaje y de patrones eficaces de aprendizaje. El cuarto estudio, analiza cómo se pueden aplicar técnicas de EMD de Multilevel (text mining-TM-) al análisis de protocolos de pensar en voz alta (think-aloud protocols -TAP-) y en el quinto trabajo, estudia técnicas de extracción de datos en las LMS, estructuración y depuración de los mismos desde el uso de técnicas de aprendizaje automático

(Machine Learning) aplicadas a la detección temprana de alumnos en riesgo de fracaso o abandono.

Área temática: Psicología de la educación y de la instrucción para nuevos entornos de aprendizaje.

Palabras clave: aprendizaje individualizado; inteligencia artificial; minería de datos; aprendizaje automático; sistemas de gestión del aprendizaje.

Symposium

TITLE OF SYMPOSIUM: *Personalized Learning in Higher Education: An Application of data mining and text-based techniques*

Coordinator: Dr. María Consuelo Sáiz Manzanares

e-mail: mcsmanzanares@ubu.es

Institution: Department of Health Sciences. University of Burgos

GENERAL ABSTRACT OF SYMPOSIUM

Recent studies in Higher Education point out that teaching methodologies must be directed at more personalized forms of interaction between student and teacher. Learning environments have changed in 21st century society. Nowadays, a high percentage of teaching, whether face-to-face or virtual, is carried out through a Learning Management System (LMS) such as Moodle. An LMS facilitates personalized learning and cultivates respect for the pace of student learning, which in turn enhances the development of Self-Regulated Learning (SRL) and motivation. All of these changes lead to effective learning and fewer students abandoning their studies. Recent studies also indicate that personalized follow-up of student-learning processes by the teacher can predict student learning outcomes by 61.3% and their behavioral patterns when using an LMS by 56.1%. Likewise, the use of LMS increases the employment of metacognitive skills, which is a relevant aspect as these skills explain 40% of learning outcomes. Nevertheless, the information that is registered in an LMS can become overwhelming. Hence, the need for Educational Data Mining (EDM) techniques in its analysis. EDM can be used in supervised machine learning (classification or prediction) and/or unsupervised machine learning (clustering). In this Symposium, the results of the application of EDM are presented in five studies with university students. In the first, the variable "collaborative work in LMS" and its effects on learning outcomes and on behavioral patterns of learning among university students are studied in two modalities of Blended learning: Supplemental Blend (SB) vs. Replacement Blend (RB). In the second work, the predictive value of clusters is studied in relation to learning outcomes and effective learning patterns. The third paper sets out an analysis of whether the prior knowledge and the metacognitive skills of students are predictors of learning outcomes and effective learning patterns. The subject of the fourth study is Multilevel Text Mining (TM) and the application of these techniques to Think-Aloud Protocols (TAP). Finally, in the fifth, data extraction, structuring, and debugging techniques in an LMS are studied through Machine-Learning techniques for early detection of students at risk of failure or likely to drop out of their course.

Thematic area: Educational Psychology and learning to work in new learning environments.

Keywords: personalized learning; artificial intelligence; data mining; machine learning; learning management systems.

Comunicación 1

Análisis del trabajo colaborativo en contextos B-Learning en Educación Superior: aplicación de EDM

*Sáiz Manzanares, M.C., **Arnaiz-González, Á., **Marticorena-Sánchez, R., **Díez-Pastor, J.F., **García-Osorio, C.I., *Escolar Llamazares, María del Camino.

**Departamento de Ciencias de la Salud, Universidad de Burgos, Burgos, España*

***Departamento de Ingeniería Civil, Universidad de Burgos, Burgos, España*

E-mail: mcsmanzanares@ubu.es

Resumen

La sociedad del conocimiento en la era del siglo XXI exige que los estudiantes desarrollen competencias de trabajo colaborativo en entornos físicos y virtuales. Los objetivos de este estudio fueron: 1) conocer si existían diferencias significativas en los grupos colaborativos respecto de los resultados de aprendizaje y de los patrones de conducta empleados en Moodle atendiendo al tipo de Blended Learning (B-Learning) utilizado (Replacement Blend vs. Supplemental Blend); 2) estudiar si el tamaño del grupo colaborativo influía en los resultados de aprendizaje y en los patrones de conducta en la plataforma de los estudiantes. Se trabajó con una muestra de 119 estudiantes universitarios de la rama de Ciencias de la Salud, 81 en el Grado en Terapia Ocupacional, y 38 en el Grado en Enfermería, en ambos se aplicaron las dos modalidades de B-Learning RB vs. SB. En el grupo RB se trabajó con 16 grupos pequeños y en el SB con 14. La metodología instruccional utilizada en ambos grupos (GRB y GSB) fue Aprendizaje Basado en Proyectos (ABP). Si bien, en GSB la instrucción se realizó utilizando en un 80% un *feedback face to face* y un 20% *feedback* a través de Moodle y en GRB el *feedback* se realizó en 80% del *feedback* en Moodle y sólo un 20% *face to face*. Se encontraron diferencias significativas en los resultados de aprendizaje y en las conductas de aprendizaje de los estudiantes en la plataforma a favor del GRB en ambos casos. Respecto del segundo objetivo en el GRB se hallaron diferencias significativas en los resultados de aprendizaje y en las conductas de aprendizaje desarrolladas en la plataforma, estas diferencias no se encontraron en el GSB. Por ello, parece ser que el tipo de metodología B-Learning utilizada incide en la dinámica de los grupos colaborativos en la plataforma. Asimismo, el uso de un RB produce una homogeneización de los resultados efectivos de aprendizaje. Los mejores resultados de aprendizaje y los patrones conductuales más efectivos en Moodle se encontraron en GRB. Futuras investigaciones irán dirigidas a estudiar las variables que potencian o inhiben el aprendizaje en entornos interactivos.

Palabras clave: aprendizaje basado en proyectos; aprendizaje mixto; minería de datos; trabajo colaborativo

Comunicación 1:**Analysis of collaborative work in B-Learning contexts in Higher Education: An application of EDM**

*Sáiz Manzanares, M.C., **Arnaiz-González, Á., **Marticorena-Sánchez, R., **Díez-Pastor, J.F., **García-Osorio, C.I., *Escolar Llamazares, María del Camino.

* *Department of Health Sciences, University of Burgos, Burgos, Spain*

** *Department of Civil Engineering, University of Burgos, Burgos, Spain*

E-mail: mcsmanzanares@ubu.es

Abstract

The knowledge society of the 21st century requires students to develop competencies in collaborative work both in face-to-face and in virtual environments. The aims of this study were: 1) To establish whether there will be significant differences in collaborative group work with regard to learning outcomes and patterns of behavior on Moodle platforms according to the type of Blended learning (B-Learning) in use (Supplemental Blend vs. Replacement Blend); and, 2) To study whether the size of the collaborative group will influence learning outcomes and patterns of behavior on the platform. We worked with a sample of 119 university students from the area of Health Sciences: 81 following the Degree in Occupational Therapy, and 38 following the Degree in Nursing. Both modalities of B-Learning (RB vs. SB) were applied to the two sub-samples. There were 16 small groups in the SR group and 14 in the SB group. In both groups (S RBG and BG) Project-Based Learning (PBL) was used. Although 80% of feedback was face-to-face and 20% was through the Moodle platform in the SBG, in contrast, 80% of feedback was through the Moodle platform and only 20% was face-to-face in the RBG. Significant differences in the learning outcomes and in the learning behaviors in favor of the RBG were found in both cases. With regard to the second objective, significant differences in learning outcomes and in learning behaviors on the Moodle platform were found in the SBG. No significant differences in either the learning outcomes or the behavioral patterns on the Moodle platform were found in the RBG. It therefore appears that the type of B-Learning methodology in use affected the collaborative group dynamics on the Moodle platform. The RB methodology homogenizes the functioning of the groups. Thus, the best learning outcomes and the most effective behavioral patterns on the Moodle platform were found in RBG. Future research will be aimed at studying which variables either facilitate or inhibit learning in interactive environments.

Keywords: project-based learning; blended learning; data mining; collaborative work.

Comunicación 2

Estudio del rendimiento académico y de las conductas de aprendizaje en LSM: aplicación de un análisis de clústeres

*Sáiz Manzanares, M.C., **Arnaiz-González, Á., *Escolar Llamazares, María del Camino.,

**Marticorena-Sánchez, R., **Díez-Pastor, J.F., **García-Osorio, C.I.

**Departamento de Ciencias de la Salud, Universidad de Burgos, Burgos, España*

***Departamento de Ingeniería Civil, Universidad de Burgos, Burgos, España*

E-mail: mcsmanzanares@ubu.es

Resumen

La docencia en la sociedad del siglo XXI se realiza en entornos semipresenciales o Blended Learning (B-Learning). La nueva sociedad del conocimiento exige cada vez más la adquisición y el uso de competencias digitales. Por ello es referencial estudiar si las competencias digitales que se han mostrado efectivas en estos entornos (consultas a la información, frecuencia de acceso al *feedback* del docente y la media de visitas por día) son predictores de los resultados de aprendizaje. Se trabajó con 119 estudiantes de tercer curso en titulaciones de ciencias de la salud. Los objetivos de este estudio fueron 1) Comprobar si el clúster de agrupación predice los resultados de aprendizaje en estudiantes en entornos RB vs. SB (Replacement Blend vs. Supplemental Blend). 2) Comprobar si el clúster de agrupación predice los patrones efectivos de interacción en la plataforma en entornos RB vs. SB. En ambos entornos RB y SB se encontraron tres clústeres de agrupación que explicaron el 66.9% y el 56.6%, respectivamente, de la varianza de los resultados de aprendizaje y de los patrones conductuales de aprendizaje. Asimismo, en los dos entornos RB y SB se hallaron diferencias significativas entre los tres clústeres de agrupación en las conductas de aprendizaje en el LMS referidas a las consultas al *feedback* realizado por el docente y en la media de vistas por día. Además, en el entorno RB se encontraron diferencias significativas en las consultas a la información complementaria. También, se realizó un Análisis de Regresión Múltiple (ARM) para comprobar si la variable pertenencia a un grupo predecía los resultados de aprendizaje. Se halló que en entornos SB lo predecía en un 87.6% y en entornos RB en un 5.9%. Futuras investigaciones irán dirigidas a analizar que variables explican las diferencias de predicción dependiendo del entorno B-Learning utilizado.

Palabras clave: análisis de clúster; análisis de regresión múltiple; supplemental blend; replacement blend.

Communication 2

Study of learning outcomes and learning behavior in LMS: An application of cluster analysis

*Sáiz Manzanares, M.C., **Arnaiz-González, Á., *Escolar Llamazares, María del Camino.,

**Díez-Pastor, J.F., **García-Osorio, C.I., **Marticorena-Sánchez, R.

* *Department of Health Sciences, University of Burgos, Burgos, Spain*

** *Department of Civil Engineering, University of Burgos, Burgos, Spain*

E-mail: mcsmanzanares@ubu.es

Abstract

Teaching in the 21st century society takes place in semi-presential or Blended-learning (B-Learning) environments. The new knowledge society increasingly demands acquisition and use of digital competences for learning. Thus, it is pertinent to study whether effective digital competences that have proven themselves to be effective in those environments (information queries, frequency of access to teacher feedback and average visits per day) will be predictors of learning outcomes. We worked with 119 third-year students studying Health Sciences degrees. The aims of this study were: 1) To confirm whether the cluster groups predict the learning outcomes in students in RB vs. SB (Replacement Blend vs. Supplemental Blend) environments. 2) To confirm whether the cluster groups predict the effective patterns of interaction in the platform in RB vs. SB environments. Three clusters in both environments (RB and SB) were found. They explain 66.9% and 56.6% of the variance in: learning outcomes and behavioral learning patterns, respectively. Likewise, significant differences in LMS learning behaviors were found in both the RB and the SB environments between these three clusters, referring to feedback to queries by teachers and to average views per day. Significant differences were also found in queries of complementary information in the RB environment. Multiple Regression Analysis (MRA) was also performed to verify whether the variable ‘belonging to one of the two groups’ predicted the learning outcomes. The variable predicted 87.6% of learning outcomes in the SB environments and 5.9% of learning outcomes in the RB environments. Future research will be aimed at analyzing why different predictions exist depending on the B-learning environment that is used.

Keywords: cluster analysis; multiple regression analysis; supplemental blend, replacement blend.

Comunicación 3

¿Son los conocimientos previos y las estrategias metacognitivas predictores de aprendizajes eficaces?

*Sáiz Manzanares, M.C., **Arnaiz-González, Á., *Escolar Llamazares, María del Camino.,

**Díez-Pastor, J.F., **García-Osorio, C.I., **Marticorena-Sánchez, R.

**Departamento de Ciencias de la Salud, Universidad de Burgos, Burgos, España*

***Departamento de Ingeniería Civil, Universidad de Burgos, Burgos, España*

E-mail: mcsmanzanares@ubu.es

Resumen

Recientes investigaciones indican que las estrategias metacognitivas (EM) explican entre el 40-52% de la varianza en los resultados de aprendizaje. Por ello, la potenciación de dichas estrategias junto con el análisis de los conocimientos previos de los alumnos previsiblemente va a facilitar una enseñanza significativa. Todo lo cual, mejorará los resultados y la calidad del aprendizaje. Además, los resultados de la investigación actual revelan que el tipo de Blended Learning (B-Learning) utilizado influye en la profundidad del aprendizaje de los estudiantes. Los entornos Replacement Blend (RB) parecen ser más efectivos que los Supplemental Blend (SB), ya que los primeros potencian el uso de las EM. Los objetivos de este estudio fueron: 1) comprobar si las EM predicen los resultados de aprendizaje en los estudiantes y si existen diferencias dependiendo del entorno B-Learning utilizado (RB vs. SB), 2) comprobar si los conocimientos previos de los estudiantes predicen los resultados de aprendizaje y si existen diferencias dependiendo del entorno B-Learning empleado (RB vs. SB). Se trabajó con una muestra de 121 estudiantes universitarios de tercer curso de la rama de Ciencias de la Salud, 58 trabajaron en un entorno SB y 63 en entorno RB. Antes de iniciar la instrucción se aplicó en ambos grupos la escala ACRAr, para medir las EM, y el Cuestionario de conocimientos previos (CCP). Una vez finalizado el semestre se realizó un Análisis de Regresión Múltiple (ARM) en cada entorno B-Learning (RB vs. SB) para comprobar el grado de predicción de las dos variables. En los entornos SB se encontró que los conocimientos previos (CP) tuvieron un valor predictivo sobre los resultados de aprendizaje de un 17.1% y las EM de un 34.3%. Asimismo, las EM que tuvieron un porcentaje mayor de predicción fueron las EM de Planificación. De otro lado, en el entorno RB los CP tuvieron un valor predictivo de 1.6% y no se observó porcentaje de predicción de las EM sobre los resultados de aprendizaje. Por ello, parece que el tipo de B-Learning condiciona el uso de los CP y de las EM en los estudiantes, lo que confirma que los entornos RB facilitan que no existan diferencias significativas en los resultados de aprendizaje de los estudiantes. Futuras investigaciones irán dirigidas a comprobar si existen diferencias antes-después de la instrucción entre estudiantes de ambos entornos en la utilización de las EM y en el grado de adquisición de los conocimientos relacionados con la materia.

Palabras clave: estrategias metacognitivas; conocimientos previos; aprendizaje mixto; análisis de regresión múltiple.

Communication 3

Are previous knowledge and metacognitive strategies predictors of effective learning?

*Sáiz Manzanares, M.C., *Escolar Llamazares, María del Camino., **Arnaiz-González, A.,

**Díez-Pastor, J.F., **García-Osorio, C.I., **Marticorena-Sánchez, R.

* *Department of Health Sciences, University of Burgos, Burgos, Spain*

** *Department of Civil Engineering, University of Burgos, Burgos, Spain*

E-mail: mcsmanzanares@ubu.es

Abstract

Recent research indicates that Metacognitive Skills (MS) can explain the variance in learning outcomes by between 40% to 52%. The strengthening of those skills together with the analysis of the prior knowledge of students will facilitate significant instruction. Likewise, this circumstance will foreseeably improve the results and the quality of learning. Moreover, the results of the present research suggest that the type of Blended learning (B-Learning) used influences in the depth of students' learning outcomes. Replacement Blend (RB) environments appear to be more effective than the Supplemental Blend (SB), as the former enhance the use of MS. The aims of this study were: 1) to confirm whether MS predict student learning outcomes and whether there are differences depending on the B-Learning environment (RB vs. SB) in use; and, 2) to confirm whether the previous knowledge of students predicts the learning outcomes and whether there are differences depending on the B-Learning environment (RB vs. SB). We worked with 121 university students in the third year of their degree in Health Sciences: 58 in an SB environment and 63 in an RB environment. Before the instruction began, the MS of the participants was measured with the ACRAr scales and a previous knowledge questionnaire (PKQ) was administered to both groups. At the end of the semester, a Multiple Regression Analysis (MRA) was performed in both environments (RB vs. SB) to check the degree of prediction of both variables. Prior knowledge (PK) in both the SB and the MS environments was found to have predictive values of learning outcomes of 17.1% and 34.3%, respectively. Moreover, the MS of planning had a higher percentage prediction of learning outcomes. In contrast, PK in the RB environment had a predictive value of 1.6% and no percentage prediction of MS on learning outcomes was observed, which confirms that the RB environments facilitate that there are no significant differences in student learning outcomes. It therefore appears that the type of B-Learning used in LMS conditions the use of both PK and MS among students. Future research will seek to verify whether there will be differences in the degree to which the students learn the subject matter before and after instruction in MS in both environments.

Keywords: metacognitive skills; prior knowledge; Blended learning; multiple regression analysis.

Comunicación 4

Un método automático para la medición del uso de estrategias metacognitivas en análisis de protocolos

**Díez-Pastor, J.F., **Arnaiz-González, Á., **García-Osorio, C.I., **Marticorena-Sánchez, R.,
*Sáiz Manzanares, M.C., *Escolar Llamazares, María del Camino.

**Departamento de Ciencias de la Salud, Universidad de Burgos, Burgos, España*

***Departamento de Ingeniería Civil, Universidad de Burgos, Burgos, España*

E-mail: mcsmanzanares@ubu.es

Resumen

Tradicionalmente la asignatura de Física está asociada a un menor rendimiento tanto en etapas universitarias como en las de Educación Secundaria y Bachillerato. Por ello, el estudio del proceso de enseñanza-aprendizaje de esta materia es referencial. El uso de las estrategias metacognitivas (EM) es uno de los predictores más significativos del éxito en el aprendizaje de los estudiantes y explica entre 40% al 52% de la varianza. Las EM se puede clasificar en EM de: orientación, planificación, evaluación y elaboración. Una de las técnicas más efectivas para evaluar el uso de dichas estrategias es el “análisis de protocolos” sobre los diálogos que se producen al “pensar en voz alta” durante la resolución de problemas. Este tipo de metodología de observación y de análisis va a permitir al docente ajustar la práctica curricular a las características de aprendizaje de sus alumnos, lo que favorece el incremento en los resultados de aprendizaje y el desarrollo de aprendizajes más profundos. No obstante, la técnica de “análisis de protocolos” es un procedimiento laborioso que exige el registro de los diálogos producidos y su clasificación de acuerdo con el tipo de EM utilizadas, tanto en su frecuencia como en el grado de profundidad. Por ello, este proceso requiere de mucho tiempo y de conocimientos especializados, lo que limita a los profesionales de la educación la aplicación de dicha técnica. Los objetivos de este estudio fueron 1) realizar una automatización de los procesos de registro y de transferencia de la información de los protocolos de pensar en voz alta; 2) aplicar un método automático basado en algoritmos de procesamiento del lenguaje natural y minería de datos que permita clasificar los diálogos de acuerdo con las EM utilizadas por el estudiante. Se analizaron 499 protocolos de pensar en voz alta recogidos a lo largo de un curso académico (28 semanas) en cuarto de Educación Secundaria Obligatoria en la asignatura de Física. El método propuesto se ha validado mediante un estudio experimental que engloba los algoritmos más populares de minería de textos. Los resultados indican un logro en la automatización de los procesos de registro, clasificación y de análisis de la información. Por ello, la aplicación de estas técnicas al análisis de la interacción (estudiante-estudiante y estudiante-docente) lleva a resultados prometedores en la facilitación del estudio y la posterior toma de decisiones para la mejora de la instrucción educativa.

Palabras clave: estrategias metacognitivas; aprendizaje automático; resolución de problemas; análisis de protocolos.

Comunication 4

An automatic method for measuring the use of meta-cognitive strategies in protocol analysis

****Diez-Pastor, J.F., **Arnaiz-González, Á., **García-Osorio, C.I., **Marticorena-Sánchez, R., *Sáiz Manzanares, M.C., *Escolar Llamazares, María del Camino.**

** Department of Health Sciences, University of Burgos, Burgos, Spain*

*** Department of Civil Engineering, University of Burgos, Burgos, Spain*

E-mail: mcsmanzanares@ubu.es

Abstract

Abstract physics concepts have always been associated with learning difficulties at the educational stages of both university and secondary education. The study of teaching-learning processes in this subject area is therefore very relevant. The use of Metacognitive Skills (MS) is one of the most significant predictors of success in student learning, as it explains between 40% and 52% of the variance. MS can be classified in terms of: reorientation, planning, evaluation, and elaboration. One of the most effective techniques for evaluating the use of such strategies is the technique of "protocol analysis" of the dialogues that occur when "thinking aloud" during problem-solving. This type of methodology of observation and analysis will allow teachers to adapt curricular practice to the learning characteristics of their students, which favors better learning outcomes and greater in-depth learning. However, the technique of "protocol analysis" is a tedious procedure in which dialogues are recorded, so as to classify them according to the type of MS that each student employs, its frequency, and depth of meaning. This process therefore requires a lot of time and specialized knowledge, which limits its application to education professionals. The aims of this study are: 1) to automate the recording and information transference processes of the protocols of thinking aloud; 2) to apply an automatic method based on natural language processing algorithms and data mining techniques that can classify the dialogues according to the MS used by each student. An analysis was conducted of 499 thinking aloud protocols, collected in the physics classroom over 28 weeks, in the fourth year of Secondary Education. The proposed method was validated through an experimental study that combines well known text-mining algorithms. The results pointed to an achievement in the automation of the processes of registration, classification, and analysis of the information. Hence, the application of these techniques to the analysis of the (student-student, student-teacher) interaction points to promising results for assisting the study of educational instruction and subsequent decision-making for its improvement.

Keywords: metacognitive skills; machine learning; problem solving; protocol analysis

Comunicación 5

Predicción de estudiantes en riesgo: aplicación de técnicas de aprendizaje automático

**Arnaiz-González, Á., **Marticorena-Sánchez, R., **García-Osorio, C. I., **Díez-Pastor., J. F., *Sáiz Manzanares, M. C., **, *Escolar Llamazares, M. C.

* *Department of Health Sciences, University of Burgos, Burgos, Spain*

** *Department of Civil Engineering, University of Burgos, Burgos, Spain*

E-mail: mcsmanzanares@ubu.es

Resumen

Uno de los grandes retos a los que se enfrentan las universidades actualmente es reducir la tasa de abandono en grados y másteres. Este abandono entre los alumnos en riesgo tiene diversas causas, pero una de las más preocupantes es el fracaso en las calificaciones obtenidas. Diversas herramientas y metodologías se vienen aplicando en las aulas con la intención de disminuir el peso de las pruebas escritas, de carácter final, en favor de una evaluación más continua. Ejemplos de estas prácticas de innovación docente pueden ser la utilización de metodologías activas como el aprendizaje basado en proyectos (ABP) o el uso de plataformas de gestión de aprendizaje (Learning Management Systems -LMS-). Los LMS son interesantes, principalmente, por dos motivos: en primer lugar, porque permiten avanzar hacia el Blended-Learning (B-Learning), (presencial y virtual), y en segundo lugar porque posibilitan el registro preciso de las interacciones de los usuarios en la plataforma, facilitando su posterior análisis. Los objetivos de este estudio fueron: 1) automatizar la extracción de la información de uso de los LMS; 2) aplicar técnicas de Aprendizaje Automático para predecir el futuro desempeño de los alumnos en las pruebas de evaluación individual. Se trabajó con 124 estudiantes del Grado en Terapia Ocupacional (83) y del Grado en Enfermería (41) de la Universidad de Burgos. A partir de datos generales de los alumnos y de sus registros de actividad, recogidos del uso de los LMS, se entrenaron sencillas técnicas de Aprendizaje Automático que permitieron predecir con una buena precisión los resultados en las pruebas de evaluación individual. Todo lo cual, facilitó la identificación con anterioridad al desarrollo de la prueba de evaluación de los alumnos en riesgo de fracaso. Este es un hito esencial para la toma de decisiones del docente que va a permitir una mejora de la práctica docente y de los resultados de aprendizaje en los estudiantes. Futuros estudios irán orientados a desarrollar técnicas de Aprendizaje Automático aplicadas a distintas materias, ya que el análisis de los registros de la actividad de los usuarios en los LMS parece ser una potente herramienta en la predicción temprana del fracaso de los estudiantes.

Palabras clave: Aprendizaje Basado en Proyectos; Sistemas de Gestión del aprendizaje; aprendizaje automático; estudiante en riesgo.

Communication 5

Prediction of drop-out in at-risk students: An application of automatic learning techniques

**Arnaiz-González, Á., **Marticorena-Sánchez, R., **García-Osorio, C.I., **Diez-Pastor, J.F., *Sáiz Manzanares, M.C., *Escolar Llamazares, María del Camino.

* *Department of Health Sciences, University of Burgos, Burgos, Spain*

** *Department of Civil Engineering, University of Burgos, Burgos, Spain*

E-mail: mcsmanzanares@ubu.es

Abstract

Nowadays, one of the biggest challenges that universities face is to reduce the dropout rates on degree and master's courses. The dropout rate in at-risk students has several causes, but poor grades are one of the most worrying. Various tools and methodologies have been applied in classrooms with the intention of decreasing the weight of written tests, of a final nature, leaning towards more continuous assessment. Examples of this may be Project-Based Learning (PBL) and the use of learning management platforms or Learning Management Systems (LMS). LMS are interesting for mainly two reasons: firstly, because they move towards Blended-Learning (B-Learning) (face-to-face and virtual) learning, and secondly because they make it possible to accurately record the interactions of the users on the platform, facilitating their subsequent analysis. The aims of this study were: 1) to automate the extraction of information on the use of LMS; and, 2) to apply Machine Learning techniques to predict the future performance of students in individual assessment tests. The study involved 124 students following the Degree in Occupational Therapy (83) and the Degree in Nursing (41) at the University of Burgos. Based on the general data of the students and their activity records, collected from the use of the LMS, simple Machine Learning techniques were trained to predict results in the individual assessment tests, with a good precision. All of these points will facilitate early identification of students at risk of failure before their evaluations. This essential milestone in teacher decision-making will increase the development of best practice in teaching and student learning outcomes. Future studies will be aimed at developing Machine-Learning techniques applied to different subjects, as the analysis of user activity records in LMS appears to be a powerful tool in early prediction of student failure.

Keywords: project-based learning; learning management system; machine learning; at-risk students.