



UNIVERSIDAD
COMPLUTENSE
MADRID

Proyecto de Innovación y Mejora de la Calidad Docente

Convocatoria 2016

Nº de proyecto 83

Big data en educación: tipologías de los estudiantes a partir del estudio de las interacciones dentro del triángulo pedagógico

Adolfo Hernández Estrada

Facultad de Comercio y Turismo

Estadística e Investigación Operativa II (Métodos de Decisión)

1. Objetivos propuestos en la presentación del proyecto

Este proyecto se presentó como continuidad a los llevados a cabo en convocatorias anteriores (conv. 2006: nº 390, conv. 2007: nº 559, conv. 2011: nº 246, conv. 2014: nº 244, conv 2015: nº 379), en los que se habían desarrollado materiales para el aprendizaje autónomo del alumno, una herramienta de aprendizaje teórico, dos herramientas de evaluación y autoevaluación, un paquete de vídeos que facilitan el aprendizaje en modelos de aula de aprendizaje inverso, siempre en el entorno del aprendizaje de la Estadística, tanto en el desarrollo de asignaturas convencionales como en el desarrollo de Trabajo Fin de Grado. El objetivo del presente proyecto se planteó para avanzar un poco más allá en el escenario de aprendizaje, dando lugar a la construcción de un modelo de análisis de los datos que facilita la plataforma Moodle, desde las distintas interacciones entre el docente, el estudiante y la materia desarrollada, de forma que esta información pueda transformarse en conocimiento y su comprensión pueda ayudar a la mejora de la práctica docente de profesores de otros campos del saber.

Es necesario que los espacios universitarios sean en este sentido a la vez que colaborativos, especializados y que podamos velar de forma conjunta por la calidad en la transmisión de los contenidos, como la aplicación de los mismos, teniendo un marcado espíritu de apertura hacia nuevas fuentes de recogida de información no siempre visibles.

Los objetivos que se enumeran a continuación fueron los que se incluyeron en el proyecto.

Objetivo principal:

Mejora de las estrategias enseñanza-aprendizaje en el proceso educativo a partir de la recogida y el estudio de la información que queda almacenada en la plataforma Moodle.

Objetivos específicos:

- Obtención de los datos que proporciona la plataforma Moodle.
- Depuración de los datos.
- Construcción de modelos de estudiantes, utilizando técnicas “Big data” (modelos de agrupamiento-clasificación).
- Aplicación de los modelos a la mejora de las estrategias de aprendizaje de los alumnos según su tipología.
- Aplicación de los modelos a la mejora de la estrategia de enseñanza del docente.

2. Objetivos alcanzados

1. Obtención de los datos que proporciona la plataforma Moodle.

Los profesores que integran el PIMCD imparten asignaturas con distintos perfiles en el Departamento, por tanto, se decidió hacer una descarga masiva de los registros que aportaba Moodle.

Así la base de datos ha quedado integrada por los registros:

- Fecha y hora, para conocer cuál es el momento en que el estudiante se conecta a cada una de las asignaturas.
- Tipo de acción, que nos permite conocer qué tipo de actividades son preferidas por los estudiantes desde el punto de vista de su trabajo interactivo con ellas.
- Dirección IP, que nos facilita determinar si los estudiantes trabajan desde puestos dentro del campus o fuera de él.

Objetivo conseguido.

2. Depuración de los datos.

Una vez descargados los datos se han construido las bases depurando todos aquellos registros que no aportaban información útil. En la actualidad, tenemos bases de datos en formato SPSS de todas las asignaturas impartidas por el equipo durante el curso escolar.

Objetivo conseguido.

3. Construcción de modelos de estudiantes, utilizando técnicas "Big data" (modelos de agrupamiento-clasificación).

Los datos estadísticos nos están permitiendo definir perfiles de estudiantes. Estos registros de datos provenientes de Moodle se han complementado con los listados obtenidos de la aplicación de gestión de alumnos (GEA), sobre todo desde aquellas variables que nos han facilitado la clasificación de los perfiles de los estudiantes como:

- Si son o no estudiantes Erasmus.
- Si la asignatura la cursan de acuerdo a las distintas formas: obligatoria, optativa y libre.
- El número de matrículas que ha realizado sobre la asignatura.
- El número de convocatorias que ha consumido.

Objetivo conseguido.

4. Aplicación de los modelos a la mejora de las estrategias de aprendizaje de los alumnos según su tipología.

El análisis de datos previo nos ha permitido diseñar implementaciones metodológicas que serán evaluadas como parte del PIMCD a desarrollarse durante el curso escolar 2017/18, desde el uso de herramientas integradas en Moodle que se adapten de manera personalizada a los distintos estudiantes, haciendo el aprendizaje más eficaz.

Objetivo conseguido.

5. Aplicación de los modelos a la mejora de la estrategia de enseñanza del docente.

Hemos podido organizar dos jornadas abiertas a la comunidad educativa, donde se han compartido experiencias de éxito para movilizar distintas estrategias de enseñanza que faciliten la adaptación a esos perfiles de estudiante.

Objetivo conseguido.

3. Metodología empleada en el proyecto

El plan de trabajo siguió una cronología temporal:

1. Selección de materias. Seleccionamos aquellas materias que respondían de forma más ajustada a la diversidad presente en nuestras facultades, dado que de esta forma podemos asegurarnos mejor la adaptación a otras situaciones, facultades y materias.

2. Obtención de los datos. Dado que el volumen de datos existente fue muy amplio, se hizo un primer estudio piloto de variables –datos- pertenecientes a un único grupo. A partir de ahí se llevó a cabo la descarga masiva de datos de otros grupos.

3. Selección de los datos. Se construyeron bases de datos en formato Excel y SPSS.

Algunos de los profesores se han formado en este tiempo en nuevos programas a través de cursos realizados durante el periodo, tanto de metodologías cualitativas (MaxQDA) como cuantitativas (Python).

4. Análisis de los datos. Se utilizaron técnicas principalmente en el área de clasificación-agrupamiento.

5. Definición de tipologías de estudiantes.

Nos apoyamos en una herramienta sobre estilos de aprendizaje que los estudiantes completaron de manera virtual en los primeros días del curso (Kolb, 1985), así como en los resultados de la aplicación de técnicas cluster sobre los datos obtenidos.

6. Estudio de la relación de las tipologías de estudiantes con las diferentes variables recogidas.

A partir de los grupos obtenidos mediante el uso de técnicas cluster se estudió la existencia de un comportamiento diferente de los individuos de cada grupo con respecto a cada variable considerada.

7. Conclusiones.

Se establecieron conclusiones a partir de todos los puntos anteriores.

Tanto la metodología como los resultados se presentaron en diferentes congresos (ver sección 5).

4. Recursos humanos

El equipo de trabajo estuvo formado, por una parte, por 10 profesores pertenecientes a la Facultad de Comercio y Turismo y la Facultad de Ciencias Económicas y Empresariales. Además, se incorporaron 3 estudiantes pertenecientes a grupos impartidos por estos profesores. Se consideró que la inclusión de la visión del estudiante mejoraría este estudio en prácticamente todas sus fases. La totalidad de los profesores tuvieron sus asignaturas dadas de alta en el Campus Virtual de la UCM, lo que permitió tener una considerable base de datos de alumnos (más de mil, dado el alto número de estudiantes en los grupos de estas Facultades) así como un número alto de interacciones (millones).

Este problema presentaba todas las características asociadas a lo que en los últimos años ha dado en denominarse “Big Data” (Mayer-Schönberger y Cukier 2013). Estas técnicas tienen su origen en métodos estadísticos aplicados en bases de datos con alta dimensión. En ese sentido, la pertenencia de la mayoría de los profesores al departamento de Estadística e Investigación Operativa II (Métodos de Decisión) resultó totalmente apropiada.

Además, varios de los profesores participantes en el proyecto tenían experiencia específica en el tratamiento de datos multivariantes, tanto a nivel docente (impartición de la asignatura “Técnicas Estadísticas Multivariantes en Gestión Comercial”, de tercer curso del grado en Comercio) como investigador.

Además han formado parte del grupo dos profesores de la Universidad Internacional de La Rioja, que han aportado al grupo tanto su formación como su experiencia en el ámbito docente, habiendo ya colaborado en anteriores PIMCD con esta facultad.

5. Desarrollo de las actividades

Las fases de trabajo siguieron los objetivos expuestos anteriormente, y que consideramos se describen en el punto 2 de objetivos de manera suficiente.

Queremos destacar que, tal como se planteaba en la solicitud, se le dio una especial importancia al apartado de “Transferencia”. Así, las actividades se vieron completadas con dos eventos de difusión abierta organizados desde este PIMCD, que nos facilitaron el trabajo con especialistas del área tanto desde la Universidad Complutense como colaborando con otras universidad y centros de investigación.

Estos dos eventos llevaron el nombre genérico de “Jornadas Big-Data en Educación”. Ambas jornadas se celebraron en la Facultad de Comercio y Turismo, Av. de Filipinas, 3, 28003 Madrid, Sala Germán Bernácer (planta baja). Para la financiación de ambas jornadas se contó con la colaboración de la Facultad de Comercio y Turismo (además de invertir en ellas el 100% de la ayuda concedida a este proyecto).

La primera jornada se celebró el 14 de Noviembre de 2016 y contó con las siguientes ponencias:

1. Esteban García Cuesta, (Universidad Europea de Madrid), Big data y análisis del aprendizaje
2. Luis de la Fuente Valentín, (Universidad Internacional de La Rioja), Aplicación e investigación de Learning Analytics: pequeñas y grandes ideas
3. Pedro José Muñoz Merino, (Universidad Carlos III de Madrid), Mejorando el proceso del aprendizaje con técnicas de minería de datos educativos.

Se acompaña programa en el anexo 2. Puede encontrarse toda la documentación en la página web:

<https://www.ucm.es/estadisticaeio-2/jornada-big-data-en-educacion->

La segunda jornada se celebró el 26 de junio de 2017, y los invitados fueron profesores del PIMCD que pudieron poner en común con los asistentes sus primeros resultados, además de contar con un ponente invitado:

1. Cristóbal Romero Morales. Profesor titular del Departamento de Informática y Análisis Numérico de la Universidad de Córdoba. Miembro del grupo de investigación KDIS (Knowledge Discovery and Intelligent Systems), Applying data mining in Moodle
2. Juan Luis Peñaloza Figueroa y María Pérez Martín. Siguiendo las huellas de Moodle en la medición del rendimiento académico universitario con datos jerárquicos
3. Adolfo Hernández Estrada, María Vela, Gregorio Tirado Domínguez, Elena Martínez Rodríguez y Juan Luis Peñaloza Figueroa. Análisis de registros de Moodle para la clasificación de estudiantes.
4. María Pérez Martín, Elena Martínez Rodríguez, Elia Fernández Molina y Blanca Arteaga Martínez. Más allá de la rúbrica tradicional. Usando big data para la evaluación de competencias.

Se acompaña programa en el anexo 3. Puede encontrarse toda la documentación en la página web:

<https://www.ucm.es/estadisticaeio-2/ii-jornada-big-data-en-educacion->

Además, siguiendo con el apartado de difusión de los resultados del proyecto, durante el desarrollo de este proyecto se han presentado los primeros avances en congresos:

Peñaloza, J.L. y Pérez, M. (2017). Following the hint of Moodle in the measurement of university academic performance with hierarchical data. *ICERI2017, 10th annual International Conference of Education, Research and Innovation*. Noviembre, Sevilla (España)

Vela-Pérez, M.; Hernández, A. Tirado Domínguez, G., Martínez Rodríguez, E. y Peñaloza Figueroa, J.L. (2017). Learning analytics to classify students according to their activity in moodle. *EDULEARN17 (9th International Conference on Education and New Learning Technologies)*. Julio, Barcelona (España).

Pérez, M., Martínez, E., Arteaga, B. y Fernández, E. (2017). La evaluación de la competencia trabajo en equipo y capacidad de comunicación, a partir de un modelo de Big Data en educación. *USATIC 2017*. Junio, Zaragoza (España)

Hernández, H., Vela-Pérez, M., Tirado, G., Martínez E. & J.L. Peñaloza (2017). Big data en educación: tipologías de los estudiantes a partir del estudio de su actividad en Moodle. *Las TIC en la Enseñanza. Experiencias en la UCM*, 130-135.

6. Anexos

Anexo 1. Referencias:

Kolb, D.A. (1985) *Learning style inventory*. Boston, Mass.: McBer and Co.

Mayer-Schönberger, V. y Cukier, K. (2013). *Big Data. La revolución de los datos masivos*. Ed. Noema.

Anexo 2. Programa de la I Jornada Big DATA en EDUCación



Facultad de Comercio y Turismo
Av. de Filipinas, 3, 28003 Madrid
Sala Germán Bernácer (planta baja)



Organiza:



Proyecto* de Innovación "Big data en educación: tipologías de los estudiantes a partir del estudio de las interacciones dentro del triángulo pedagógico"

* Universidad Complutense de Madrid (UCM)
* Universidad Internacional de La Rioja (UNIR)

10:00 a 10:10

Inauguración

Jornada de INNOVACIÓN



14-Nov-2016

Big DATA en EDUCación

Inscripción:

<https://goo.gl/forms/IXeDTbSBzpD4LIAc2>

-Aforo limitado-



10:10-11:00

Esteban García Cuesta
(Universidad Europea de Madrid)
Big data y análisis del aprendizaje

11:00-11:30

Pausa para café



11:30-12:30

Luis de la Fuente Valentín
(Universidad Internacional de La Rioja)
Aplicación e investigación de Learning
Analytics: pequeñas y grandes ideas

12:30-13:30

Pedro José Muñoz Merino
(Universidad Carlos III de Madrid)
Mejorando el proceso del aprendizaje con
técnicas de minería de datos educativos

13:30-14:00

Preguntas y conclusiones

II Jornada de INNOVación



Big DATA en EDUCación

Siguiendo las huellas de Moodle

Organiza:

Proyecto* de Innovación **“Big data en educación: tipologías de los estudiantes a partir del estudio de las interacciones dentro del triángulo pedagógico”**

* Universidad Complutense de Madrid (UCM)
* Universidad Internacional de La Rioja (UNIR)

Los planteamientos actuales en el entorno universitario, teniendo en cuenta la incorporación de tecnologías, construyen un escenario donde los datos procedentes de las interacciones dentro del triángulo didáctico, saber-profesor-alumno (Houssaye, 1988) han aumentado considerablemente tanto en número como en posibilidades, y requieren una intervención por parte de los docentes que puedan integrarlos en la mejora de su práctica docente y por lo tanto, al incremento de resultados favorables en el aprendizaje significativo de los estudiantes para su incorporación al mercado laboral.

La plataforma virtual Moodle facilita un entorno dinámico para utilizarlo como innovación en el proceso de enseñanza-aprendizaje e incorporarlo a **prácticas de investigación-acción**. Este hecho se justifica en el aporte de una gran cantidad de información procedente de distintas fuentes (accesos del alumno, materiales, tipología de los dispositivos, formación de grupos, etc.), que pueden suministrar datos que analizados de forma conveniente lleguen a definir modelos de actuación que ayuden tanto al profesor en ejercicio como al profesor novel que comienza su experiencia docente en el departamento.

Como medio de colaboración de esta investigación se organiza esta jornada donde contaremos con **ponentes de amplia experiencia docente e investigadora, y donde se aportarán los primeros resultados de este proyecto de investigación**. Consideramos el espacio universitario un marco de colaboración especializado, donde de forma conjunta podamos construir espacios que permitan el trabajo conjunto de recogida y análisis de un enorme volumen de datos, cuyo fin principal es la mejora del aprendizaje.

26 DE JUNIO DE 2017

Inauguración 

10:00-10:10 h.

II

Jornada de INNOVación



Big DATA en EDUCación



Applying Data Mining in Moodle



Cristóbal Romero Morales. Profesor titular del Departamento de Informática y Análisis Numérico de la Universidad de Córdoba. Miembro del grupo de investigación KDIS (Knowledge Discovery and Intelligent Systems).

Pausa-café 

11:30-12:00 h.

Siguiendo las huellas de Moodle en la medición del rendimiento académico universitario con datos jerárquicos

Juan Luis Peñaloza Figueroa y María Pérez Martín
Universidad Complutense de Madrid



Siguiendo las Huellas de Moodle en la Medición del Rendimiento Académico Universitario con Datos Jerárquicos

Juan Luis Peñaloza Figueroa y María Pérez Martín
Departamento de Estadística e Investigación Operativa II (Métodos de Decisión)

Análisis de registros de Moodle para la clasificación de estudiantes

Adolfo Hernández Estrada, María Vela Pérez, Gregorio Tirado Domínguez, Elena Martínez Rodríguez, Juan Luis Peñaloza Figueroa
Universidad Complutense de Madrid



Análisis de registros de Moodle para la clasificación de estudiantes

Adolfo Hernández, Elena Martínez, Juan Luis Peñaloza, Gregorio Tirado, María Vela
Departamento de Estadística e Investigación Operativa II (Métodos de Decisión)

Más allá de la rúbrica tradicional. Usando big data para la evaluación de competencias

María Pérez Martín, Elena Martínez Rodríguez, Elena Fernández Molins y Blanca Arceaga Martínez
Universidad Complutense de Madrid y Universidad Internacional de La Rioja



MÁS ALLÁ DE LA RÚBRICA TRADICIONAL. USANDO BIG DATA PARA LA EVALUACIÓN DE COMPETENCIAS

unir