



PATRONES DE COMPORTAMIENTO EN EL USO DE LAS AULAS VIRTUALES DE LA UNIVERSIDAD NACIONAL DEL ALTIPLANO ÁREA DE INGENIERÍAS UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS

PATTERNS OF BEHAVIOR IN THE USE OF VIRTUAL CLASSROOMS OF THE NATIONAL UNIVERSITY OF THE ALTIPLANO ENGINEERING AREA USING DATA MINING TECHNIQUES

Teresa Paola Alvarez Rozas¹

¹Universidad Nacional del Altiplano, Escuela Profesional de Ingeniería Estadística e Informática, Av. Sesquicentenario N° 1154 Ciudad Universitaria, Puno, Perú, tpalvarez@unap.edu.pe

RESUMEN

La investigación se orienta en la búsqueda del comportamiento, e identificación de patrones desconocidos en una plataforma virtual (Moodle), haciendo uso de big data, de los estudiantes de Ingenierías de la Universidad Nacional del Altiplano II semestre año 2017, considerando para ello como objetivos específicos analizar el uso de la plataforma virtual, establecer asociaciones de las variables e identificar patrones desconocidos en la plataforma virtual. Obteniendo los siguientes resultados: de las 16 escuelas de ingenierías presentan mayor actividad Ingeniería de Minas, Ingeniería Mecánica Eléctrica, Ingeniería de Sistemas e Ingeniería Económica, también sobresalen el uso de los siguientes módulos: course (curso), assign (asignar), resource (recurso), quiz (examen) y fórum (foros), a horas:

18, 19 y 20. Al relacionar las variables, escuela profesional, módulos y horas, encontramos los siguientes patrones desconocidos: el módulo VISTA en las escuelas de Ing. de Minas e Ing. Topográfica y Agrimensura, presentan un comportamiento similar, con una confiabilidad del 97 %; las escuelas de Ingeniería de Minas, Agroindustrial, Económica, Electrónica y Arquitectura y Urbanismo, presentan comportamientos similares en términos de horas de uso y actividades con un nivel de confiabilidad del 98 %. Asociando con mayor precisión se obtuvo los siguientes patrones: en el cluster 0 ingeniería de Minas tiene más actividad con el módulo Course y View a las 18 horas, y en el cluster 1 con un 23 % de la data, Ing. Económica registra mayor actividad en asignación de tareas a las 22 horas.

Palabras clave: Aulas virtuales, big data, minería de datos, patrones y plataforma virtual.

ABSTRACT

The article focuses on the search for behavior, and identification of unknown patterns in a virtual platform (Moodle), making use of big data of Engineering students of the National University of the Altiplano II semester of 2017, considering for this purpose specific analyze the use of the virtual platform, establish associations of the variables and identify unknown patterns in the virtual platform. Obtaining the following results: of the 16 engineering schools have the greatest activity in Mining Engineering, Electrical Mechanical Engineering, Systems Engineering and Economic Engineering, the use of the following modules also stands out: course (course), assign (assign), resource (resource), quiz (exam) and forum (forums), at hours: 18, 19 and 20. When relating the variables, professional school, modules and hours, we find the following unknown patterns: the VISTA module in the Ing. Mines and Topographic Engineering and Surveying, have a similar behavior, with a reliability of 97 %; The schools of Mining Engineering, Agroindustrial, Economic, Electronic and Architecture and Urbanism, have similar behaviors in terms of hours of use and activities with a 98 % reliability level. Associating with greater precision the following patterns were obtained: in the cluster 0 Mining Engineering has more activity with the Course and View module at 18 hours, and in cluster 1 with 23 % of the data, Economic Engineer registers greater activity in homework assignment at 10 pm.

Keywords: Big data, data mining, patterns, virtual classrooms and virtual platform.

*Autor para correspondencia: tpalvarez@unap.edu.pe



INTRODUCCIÓN

La generación masiva de datos es un problema de infoxicación, disponemos de tanta información que a veces es imposible organizarlas con efectividad, la clave es encontrar patrones que expliquen el comportamiento de los datos (Camargo *et al.* 2014); el amplio desarrollo del mundo tecnológico y de las comunicaciones hace posible la aparición de nuevas fórmulas educativas entre las que se encuentra el e-learning, b-learning como formación basada en TIC mediante el uso de plataformas virtuales (Sinche 2016), es justamente los alumnos que ingresan a la universidad enfrentan problemas en su rendimiento académico, al trabajar con predicciones mediante minería de datos con los factores social, económico y académico (Yamao 2018); del mismo modo se probó con otros factores como cantidad de asignaturas, comedor universitario, carrera profesional y deuda con la universidad (Holgado 2018).

Para los Sistemas de Gestión de Aprendizaje se relacionó recursos utilizados, acciones realizadas, lugar, horario y mes de trabajo, identificaron patrones en torno al uso de la plataforma virtual Moodle (Samaniego *et al.* 2014); es así que se implementó un estudio cuasi experimental con post prueba, formando dos grupos uno que trabajo en salón y otro en aulas virtuales, llegando a la conclusión que con aulas virtuales mejoraron sus calificaciones (Cepeda 2017); otro factor que se trabajó en aulas virtuales es la deserción académica con minería de datos, obteniendo información de

aquellos alumnos que son propensos a abandonar la carrera (Sulla 2015). Al agrupar a los alumnos en alto, medio y bajo con técnicas no supervisadas de minería de datos utilizando los algoritmos K-means, se pudo obtener grupos de mejor calidad y grupos que necesitan reforzamiento (Ochoa *et al.* 2017); dependiendo del horario de los estudiantes a las aulas virtuales, quienes se conectan mayormente en horarios del día, su actividad de interacción mensual es equitativa excepto los meses de inicio y fin del periodo que son menores (Hidalgo 2018), por tanto podemos afirmar que la enseñanza no solo se da en forma presencial sino también en las plataformas e-learning o tutores inteligentes, y con minería de datos se extrae patrones y modelos (Garcia 2016); la metodología PACIE el cual utiliza herramientas de la plataforma Moodle, facilita el diseño del aula virtual, mejorando el rendimiento académico (Flores 2016).

En busca de mejoras en el rendimiento académico y siguiendo el enfoque OntoSemiotico, se encontró modelos, que muestran el comportamiento de los estudiantes, permitiendo al docente redirigir las aptitudes de los alumnos (Muñoz 2015). Al trabajar en aulas de educación semi presencial con ambientes virtuales, bajo una analítica de visualizaciones, en un proceso exploratorio, intuitivo, analítico, multivariado y relacional (Aristizábal 2017). El reto está en procesar grandes cantidades de datos y convertirlas en decisiones inteligentes y oportunas, la presentación de un metamodelo que captura, transporta, transforma y analiza datos (Rojas 2017), para facilitar y apoyar la

labor de los grupos encargados de trabajar en la identificación, prevención y control en los riesgos psicosociales, se comprobó que la metodología J48, es la más efectiva, apoyando la labor de los grupos encargados de trabajo (Mosquera *et al.* 2016).

Minería de datos no solo trabaja con educación también abarca otras ramas como en las empresas, descubriendo patrones en el uso de las herramientas TIC para gestionar dichos conocimientos (Marulanda *et al.* 2017); entonces la metodología que mejor se adapta a la industria de servicios utilizando herramientas como el comportamiento de patrones y predicción de tendencias es CRISP – DM (Benalcázar 2017). En cuanto a los Sistema de Gestión Bibliotecario, basado en grandes volúmenes de información, la minería de datos trabajó para mejorar y obtener patrones útiles en la solución de objetivos propuestos (Ruiz y Romero 2018); se planteó conocer patrones de predicción generando un modelo de clasificación que asigna elementos no evidenciados para estimar la necesidad de los medicamentos y mejorar su capacidad de control de provisiones (Melo 2018). Justamente en el ámbito de la salud la metodología de extracción de conocimientos mediante en el uso de las TICs (Gutiérrez 2017). En cuanto a los datos de dominio ambiental, las ventajas que nos da la tecnología asociada a big data para su tratamiento, es la relación entre la inteligencia ambiental y las variables de precipitación, temperatura y presión barométrica, para hacer predicciones (Hernández 2016); por último se desarrolló un estudio basado en técnicas de

minería de datos de una novedosa metodología que basándose en técnicas de agrupamiento, busca la existencia de ciertos patrones precursores temporales previos a la ocurrencia de terremotos de magnitud media y elevada (Florido 2015).

La investigación plantea analizar el uso de la plataforma en el rendimiento académico, que sirve como fuente de conocimiento para los docentes, así mismo establecer asociaciones entre las variables en el uso de la plataforma y finalmente identificar patrones desconocidos que nos oriente sobre el comportamiento de los datos. La investigación es de carácter cuantitativa, el diseño no experimental y el tipo de investigación correlacional.

MATERIALES Y MÉTODOS

Ámbito o lugar de estudio

El estudio se aplica en estudiantes de Ingenierías II semestre año 2017, de la Universidad Nacional del Altiplano ubicado en la ciudad de Puno, perteneciente a la Región Puno - Perú, según las coordenadas, su ubicación es de 15°50'31" de Latitud Sur y 70°01'11" de Longitud Oeste del Meridiano de Greenwich con una altitud sobre el nivel del mar de 3825 m.s.n.m.

Descripción de métodos

La Universidad Nacional del Altiplano, cuenta con 9 facultades dentro de las cuales se tiene 16 escuelas profesionales, con un total de más de 10,000 estudiantes quienes dentro del proceso de enseñanza y aprendizaje trabajan con aulas

virtuales en la plataforma Moodle. La investigación describe los siguientes objetivos:

Analizar el uso de la plataforma virtual

Diseño de la muestra

La muestra está compuesta de las actividades realizadas en las aulas virtuales por los estudiantes de Ingenierías de la Universidad Nacional del Altiplano, II semestres año 2017, expresado en 191 990 actividades, aplicando el muestreo no probabilístico por conveniencia a criterio del investigador. Muestra que es válido para los tres objetivos específicos.

Materiales, insumos y equipos utilizados

Debido a la gran cantidad de datos los cuales fueron decodificados, se trabajó con un equipo con las siguientes características: Computadora personal con microprocesador de 2.0 GHz- i5 con 4 Gb de RAM, disco duro de 2 Gb, equipo en el que se instaló el software libre WEKA, que es una herramienta para el aprendizaje automatizado y minería de datos (Córdoba Fallas, 2013), con licencia GNU, esto debido a la gran cantidad de datos, los cuales fueron decodificados para su entendimiento. Software y equipo que fueron utilizados en los 3 objetivos específicos.

Para el análisis de los datos se trabajó en forma descriptiva las variables: Escuela profesional, módulos, actividad y hora.

Asociación de las variables en la plataforma virtual

Con respecto a las variables se asoció cada una de las 16 escuelas profesionales con los 9 módulos desde las 00 horas hasta las 23 horas:

Escuela profesional: Arquitectura y Urbanismo, Ciencias Físico Matemático, Ing. de Minas, Ing. Agrícola, Ing. Agronómica, Ing. Civil, Ing. de Sistemas, Ing. Económica, Ing. Electrónica, Ing. Estadística e Informática, Ing. Geológica, Ing. Mecánica Eléctrica, Ing. Metalúrgica, Ing. Química e Ing. Topográfica y Agrimensura

Módulo de las aulas virtuales: Curso, Blog, Foro, Asignación de Tareas, Examen, Archivo, Folder, URL (Dirección WEB) y Chat.

Hora: que corresponde desde las 00 horas hasta las 23 horas.

Identificación de Patrones

Las variables que fueron asociadas para encontrar patrones de conocimiento son: Escuela profesional, modulo, actividad y hora en las que fueron trabajadas. La aplicación estadística que se trabajó es con la técnica de APRIORI, el cual define el porcentaje de confiabilidad al encontrar similitudes respecto al comportamiento de las variables.

RESULTADOS Y DISCUSIÓN

Analizar el uso de la plataforma virtual

Una vez que tenemos la data limpia, empezamos con el estudio descriptivo, analizaremos el uso de la plataforma virtual, el cual se optó por trabajar con las actividades de la plataforma virtual como variable dependiente y las escuelas de ingenierías, como variable

independiente, debido a la notable presencia que tiene esta variable. Obtenemos los siguientes resultados: Ingeniería de Minas sobresale de las demás, obteniendo 56 329 registros seguidamente de Ingeniería

Económica con 37 824 y aparece Ingeniería Electrónica, obteniendo 30 887 actividades; escuelas que realizan la mayor actividad dentro de las aulas virtuales (Figura 1).

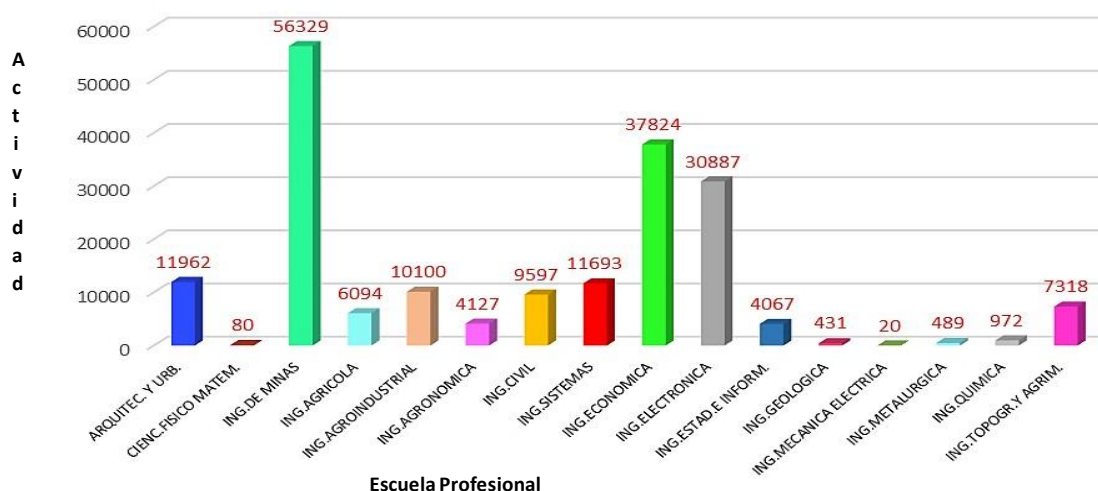


Figura 1. Actividad por Escuela Profesional. Coincidimos con el autor en indicar que para analizar las variables relacionadas se debe trabajar con un conjunto de técnicas que permite explorar la base de datos obteniéndose los resultados de forma automática (Aristizábal 2017); de la misma forma indicar que trabajando con big data que son grandes cantidades de datos, almacenando información, y trabajándolas con las variables que más actividad presentan, ayuda para la mejor toma de decisiones (Camargo *et al.* 2014).

Al trabajar con los módulos como la variable independiente y la variable dependiente actividad que se presentan en la plataforma virtual resaltan aquellas que tienen mayor presencia, obteniendo los siguientes resultados: curso con 81 940 actividades, siendo el más visitado por los estudiantes, seguido de asignación de tareas con 42 964, archivo con 20 985, examen con 19 057 y fórum con 14 510 como se presenta en la (Figura 2).

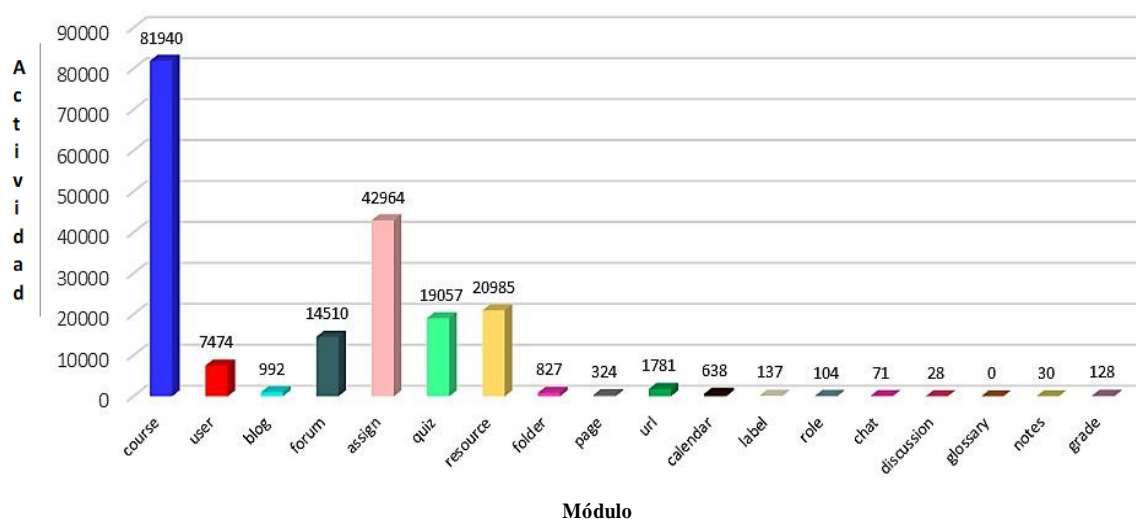


Figura 2. Actividad por Uso de Módulos.

Teniendo un amplio desarrollo del mundo tecnológico, resalta como formación básica el uso de las plataformas virtuales y las actividades que se realiza en ella (Sinche 2016); concordamos con el autor en indicar que las aulas virtuales y sus distintas actividades influyen significativamente en el aprendizaje de los estudiantes universitarios (Aguilar 2014).

Por último, dentro del estudio descriptivo se presenta las horas en las que con mayor

frecuencia ingresan los estudiantes a las aulas virtuales, datos que fueron tomados desde las 00 horas hasta las 24 horas como se presenta en la Figura 3, donde Observamos que el tiempo esta detallado según las horas de la mañana y la noche, con menor frecuencia las 3 de la mañana en 399 actividades y la hora con mayor frecuencia son las 19 horas con 16 200 actividades (Figura 3).

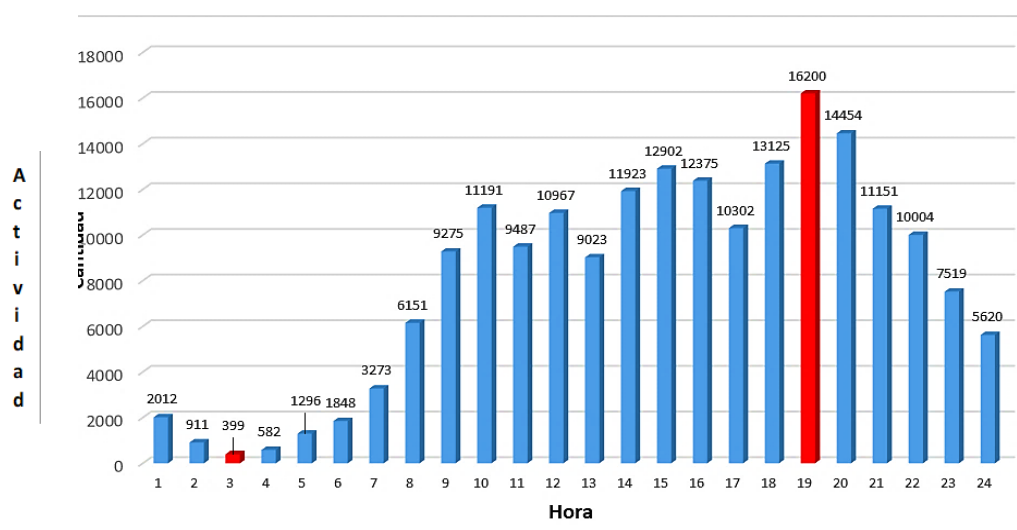


Figura 3. Actividad Tiempo de Uso de la Plataforma.

Coincidimos con los autores Samaniego *et al.* (2014) en indicar que el tiempo (hora) es otro indicador que nos ayuda a encontrar patrones de conocimiento; obteniendo para ambos estudios que la mayor fluidez en el uso de la plataforma virtual se encuentra entre las 9 de la mañana hasta las 10 de la noche (22 horas), dependiendo del horario de los estudiantes a las aulas virtuales, quienes se conectan mayormente en horarios del día, su actividad de interacción mensual es equitativa excepto los meses de inicio y fin del periodo que son menores (Hidalgo 2018), la metodología PACIE el cual utiliza herramientas de la plataforma Moodle, facilita el diseño del aula virtual, mejorando el rendimiento académico (Flores 2016).

Identificación de Patrones

Dentro de las fases de la metodología KDD resolvemos el problema refiriéndonos al paradigma apropiado de Minería de Datos para conocer la asociación de las variables y encontrar patrones que nos orienten sobre el comportamiento de los datos, analizando la relación de las variables: escuelas profesionales, módulos y horas de trabajo

La relación entre la hora de acceso al sistema, escuela profesional y la actividad realizada, siendo el mayor ingreso a las aulas virtuales a las 19 horas con 16 200 actividades, teniendo en este horario el ingreso de las escuelas de Ing. Agrícola e Ing. Metalúrgica, resaltando nuevamente las escuelas de Ing. de Sistemas, Ing. de Minas, Ing. Económica, Ing. Mecánica Eléctrica e Ing. Electrónica (Figura 4).

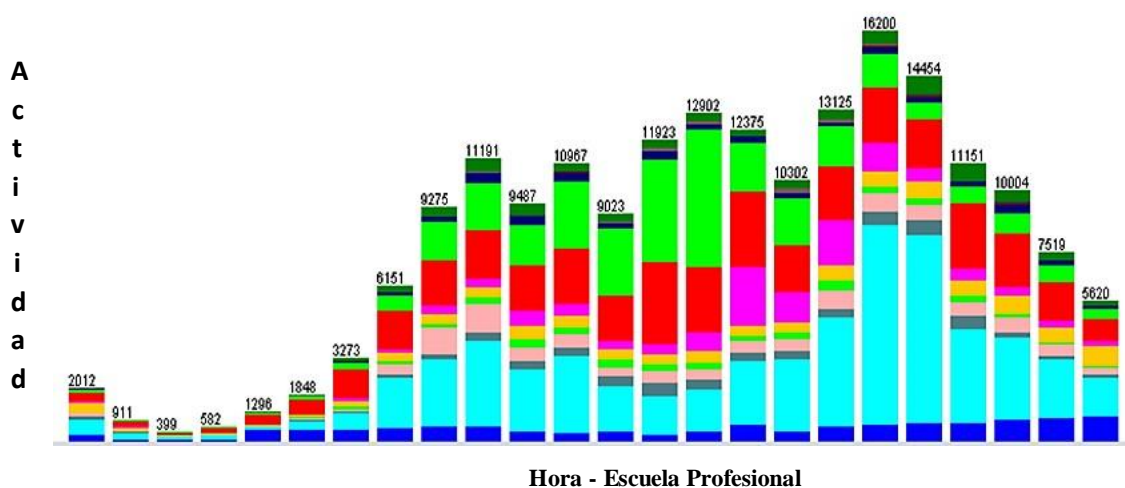


Figura 4. Horas de acceso al Sistema por Escuela Profesional

El estudio realizado coincide con Holgado (2018) al relacionar las escuelas profesionales con otras variables (cantidad de asignaturas, deuda con la universidad), encontrando

patrones como el bajo rendimiento, lo cual ayudo a las autoridades a tomar decisiones que ayuden al alumno. De la misma forma indicar que el análisis de los datos en base a

visualización, realiza un proceso exploratorio, intuitivo, analítico, multivariado y relacional que hace comprensible e interpretable el conjunto de datos (Aristizábal 2017).

Otra relación son las horas de acceso al sistema y módulos del aula virtual, siendo el mayor ingreso a las aulas virtuales las 19 horas con el módulo curso, a las 14 horas asignación de tareas, examen a las 15 y 19 horas, por último, el módulo recurso a las 18 horas (Figura 5).

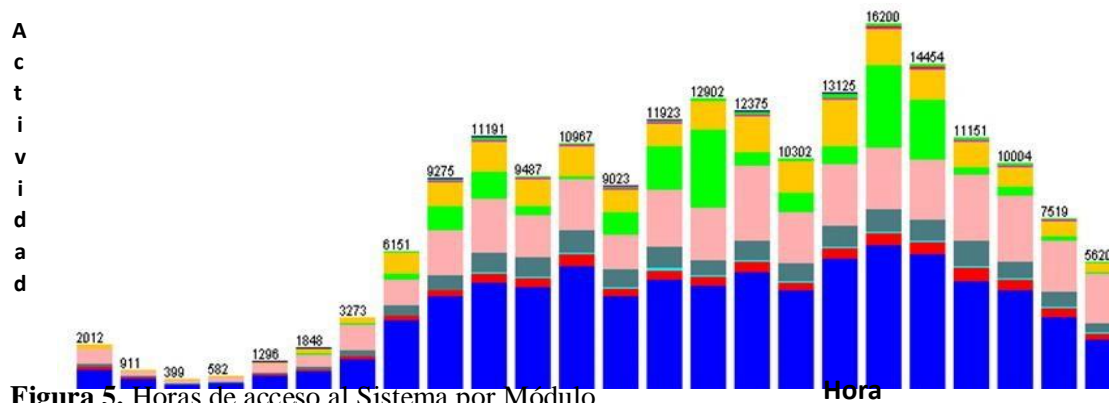


Figura 5. Horas de acceso al Sistema por Módulo

Confirmamos con el autor que al relacionar variables identificamos patrones, que son las variables que más resaltan en una asociación dentro de la plataforma virtual Moodle (Samaniego *et al.* 2014). Así mismo la asignación de tareas tiene mayor actividad, debiendo ser asistido por el docente para tener mejores resultados (Cepeda 2017).

Concluyendo con las asociaciones se trabajó con las iteraciones dentro del método K-means, donde se evidencia que al utilizar toda la data determinamos que la escuela profesional de Ingeniería de Minas tiene la mayor actividad en

la plataforma virtual, con el uso del módulo Curso(course), en la Actividad Vista(view) y la hora en que más realizan esta interacción es a las 18 horas. Con solo el 77 % (cluster 0), afirmamos el resultado anterior, pero cuando se trabaja al 23 % (cluster 1) se observa a la escuela profesional de Ing. Económica, con el modulo asignación de tarea, en la actividad vista y realizado a las 22 horas. Se realizó también pruebas con 3, 4 y 5 clúster, los cuales reafirman los resultados del análisis de clúster 1, llegando a obtener patrones de conocimiento, que ayuden en la toma de decisiones (Tabla 1).

Tabla 1. Centroides del grupo final

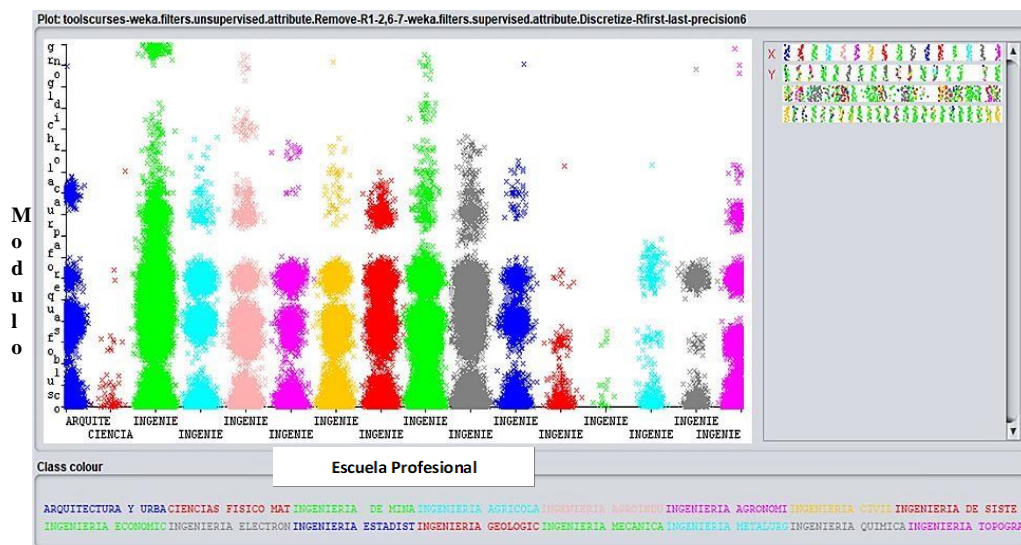
Attribute	Full Data	Cluster #	
		0	1
	(191 990,0)	(147168) 77 %	(44822,0) 23 %
Escuela	Ingeniería de Minas	Ingeniería de Minas	Ingeniería Económica
Module	Course	Course	Assign
Activity	View	View	View
hour	18	18	22

Ochoa *et al.* (2017) afirman que utilizando el algoritmo K-means obtenemos grupos de mejor calidad para conocer patrones de conocimiento. Así mismo al clasificar a los individuos en k grupos, se puede intercambiar entre los miembros de los clusters para tener una mejor partición (De la Fuentes 2014)

Representación gráfica

Es necesario tener el comportamiento de la actividad en la plataforma mediante una representación gráfica, donde las escuelas profesionales son la variable independiente y los módulos las variables dependientes.

Se observa que la escuela de Ingeniería de Minas presenta la mayor actividad resaltando de entre las otras escuelas, seguidamente de la escuela de Ingeniería Económica, Ingeniería Electrónica e Ingeniería de Sistemas. Mientras que las escuelas de Ciencias Físico Matemático, Ingeniería Geológica e Ingeniería Mecánica son en extremo pobre con su participación y/o uso de la plataforma. Se observa también a las escuelas de Ingeniería Metalúrgica e Ingeniería Química tratando de salir de este grupo (Figura 6).

**Figura 6.** Relación entre Escuelas Profesionales y los Módulos.

consideramos que hay dos actividades que se presentan en casi todas las escuelas profesionales como son adicionar o crear y visualizar contenido, reforzando esto que los puntos de color azul (add) implica que para que se tenga actividad en la plataforma tiene que crear y/o adicionar contenido, pero lo interesante está en los puntos rosados que está presente en casi todas las escuelas y esto es por la actividad visualizar contenido; implica que la actividad de hacer una simple vista de la plataforma prevalece en relación a otras actividades que sí podrían mostrar la real interactividad con el contenido por parte de los usuarios-alumnos (Figura 7).

Al relacionar dos variables, escuela profesional (variable independiente) y la actividad (variable dependiente), se muestra una importante actividad del uso de todos sus módulos por las escuelas profesionales de Ing. de Minas, Ing. Economía, Ing. Electrónica, Ing. de Sistemas que son las que más resaltan, también

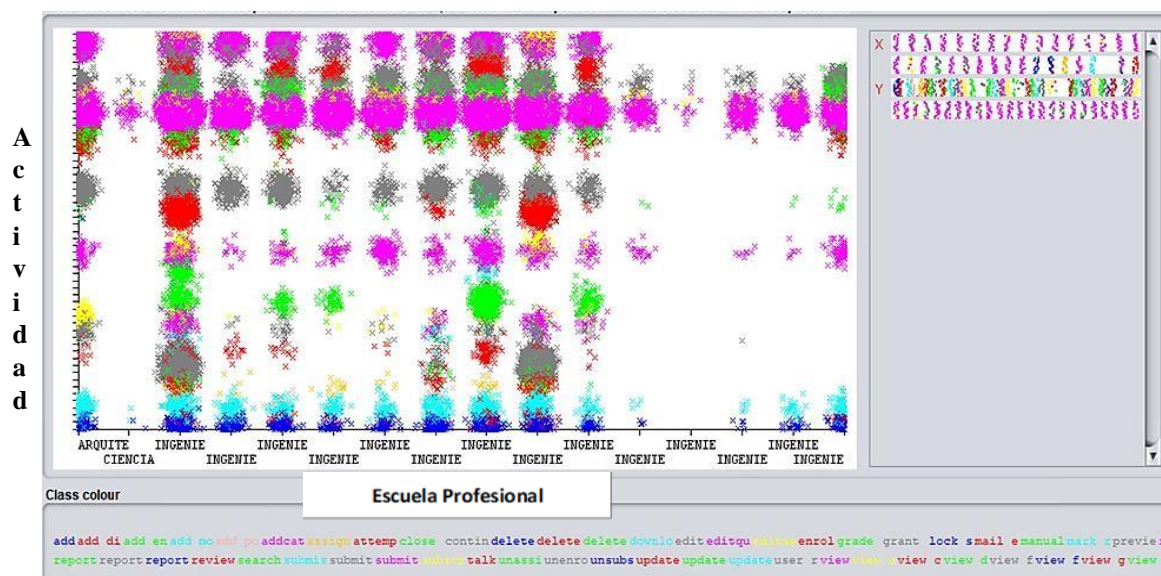


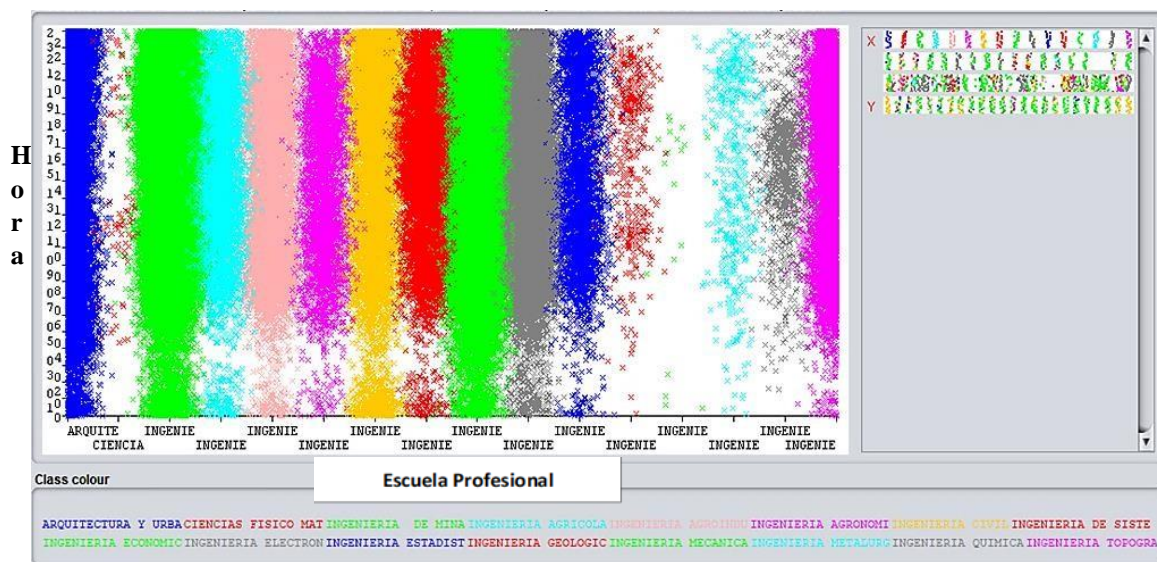
Figura 7. Relación entre las Escuelas Profesionales y la Actividad.

Según el autor que muestra en los resultados de su investigación, indica que la actividad que más registros han acumulado fue de tipo exploratorio, que ha sido la actividad ver, “ver tema de debate” y “ver foro de discusión”. El alumnado ha empleado la mayor cantidad de

tiempo en explorar la herramienta (Espigares & García 2013).

Por ultimo relacionando las variables: escuela profesional (variable independiente) y las horas (variable dependiente) dentro de la plataforma virtual, se evidencia las siguientes asociaciones:

las escuelas como Ciencias Físico-Matemáticas, Ingeniería Metalurgia, Ingeniería Geológica e Ingeniería Química continua siendo casi nula su participación a cualquier hora del uso de la plataforma virtual (Figura 8).



obteniendo las siguientes asociaciones, También coincidimos con el estudio realizado con De la Rosa (2011) donde nos indica que, en este proceso de analizar el uso de la plataforma virtual, principalmente en los módulos foro y cumplimiento de tareas, el principal impulsor es el docente profesor, dejando paso a un sistema basado en el aprendizaje, donde el alumno es responsable de su propio aprendizaje y el profesor debe buscar y analizar la metodología y los medios más adecuados que ayuden al alumno en este proceso.

En el uso de aulas virtuales de los estudiantes de Ingenierías de la Universidad Nacional del Altiplano, resaltan los módulos Vista, Foro,

Asignación de Tareas y Archivo, siendo las horas más frecuentes las 17 horas, que son patrones cuyo comportamiento estadístico es altamente significativo. Con el método de clasificación J48 dentro del análisis de minería de datos se pudo validar el proceso, llegando a determinar las siguientes asociaciones: La escuela de Ingeniería de Minas al presentar la mayor actividad obtiene una mayor asociación en los módulos: Actividad a las 00 horas FORO a horas 6:00 a.m., Asignación de Tareas a horas 20:00 p.m., examen y archivo en distintas horas, pero solo el módulo Folder aparece en esta escuela. Con la metodología KDD, implementando Minería de Datos en el programa WEKA y aplicando la técnica de

asociación A PRIORI, se pudo encontrar los siguientes patrones desconocidos: En el módulo Vista las escuela de Ingeniería de Minas e Ingeniería Topográfica y Agrimensura presentan un mismo comportamiento con una confiabilidad del 97 %, en tanto las escuelas de Ingenierías: Minas, Agroindustrial, Economía, Electrónica y la escuela de Arquitectura y Urbanismo presentan un comportamiento similar en términos de horas de uso, tiempo, uso de recursos y actividades con un nivel de confianza del 98 %.

CONFLICTO DE INTERÉS

La autora de iniciales (TPAR), no tiene conflictos de ninguna índole.

REFERENCIAS

- Aguilar M. 2014. Influencia de las Aulas Virtuales en el aprendizaje por competencias de los estudiantes del curso de internado Estomatológico de la Facultad de Odontología de la Universidad de San Martín de Porres. (Tesis de doctorado), Lima, Perú. Retrieved from http://www.repositorioacademico.usmp.edu.pe/bitstream/usmp/1069/1/aguilarm_vm.pdf
- Aristizábal J. A. 2017. Diseño y aportes de un modelo para minería de datos educativos en aulas de educación media de carácter presencial, (Tesis de Doctorado), Universidad de Santo Tomás, Bogotá, Colombia. Retrieved from <https://repository.usta.edu.co/bitstream/handle/11634/3945/Aristizabaljorge2017.pdf?sequence=1&isAllowed=y>
- Benalcázar J. B. 2017. Análisis comparativo de metodologías de minería de datos y su aplicabilidad a la Industria de Servicios. (Tesis Maestría.) Universidad de las Américas de Ecuador, Ecuador. Retrieved from <http://dspace.udla.edu.ec/bitstream/33000/7547/6/UDLA-EC-TMGSTI-2017-11.pdf>
- Camargo J.J., Camargo J.F., Joyanes L. 2014. Conociendo Big Data. Facultad de Ingeniería, vol.24 no. (ISSN 0121-1129). Retrieved from http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-11292015000100006

-
- Cepeda L. G. 2017. Estrategia metodológica del uso de aulas virtuales en el proceso de enseñanza aprendizaje universitario. (Tesis de Doctorado) Universidad Nacional Mayor de San Marcos. Lima, Perú. Retrieved from http://cybertesis.unmsm.edu.pe/bitstream/handle/cybertesis/6058/Cepeda_al.pdf?sequence=1&isAllowed=y
- Córdoba L. 2013. Minería de Datos_ WEKA. Retrieved from <http://cor-mineriadedatos.blogspot.com/2011/06/weka.html>
- De la Fuentes S. 2014. Analisis de Conglomerados. Madrid, España. Retrieved from <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/SEGMENTACION/CONGLOMERADOS/conglomerados.pdf>
- De la Rosa J. C. 2011. Aplicación de la plataforma moodle para mejorar el rendimiento académico en la enseñanza de la asignatura de cultura de la calidad total en la Facultad de Administración de la Universidad del Callao. (Tesis de maestría). Universidad Nacional de San Marcos, Lima, Perú. Retrieved from http://cybertesis.unmsm.edu.pe/bitstream/handle/cybertesis/2542/rosa_rj.pdf;jsessionid=8768E74C93A9E66F2366352BE632CB7C?sequence=1
- Flores E. F. 2016. “ Influencia de la plataforma Moodle en el rendimiento academco de los estudiantes del curso de precalculo I de la Universidad Continental. (Tesis de maestría)” Universidad Nacional del Centro del Perú, Huancayo, Perú. Retrieved from <http://repositorio.uncp.edu.pe/bitstream/handle/UNCP/3891/FloresAlvarez.pdf?sequence=1&isAllowed=y>
- Florido E. 2015. Metodología basada en minería de datos para el descubrimiento de patrones precursores de terremotos de magnitud media y elevada. (Tesis de doctorado), Universidad Pablo de Olavide, Sevilla, España. Retrieved from https://www.researchgate.net/publication/282186653_Metodologia_basada_en_mineria_de_datos_para_el_descubrimiento_de_patrones_precursores_de_terremotos_de_magnitud_media_y_elevada
- García F. J. 2013. Aplicación de técnicas de Minería de Datos a datos obtenidos por el Centro Andaluz de Medio Ambiente (CEAMA), (Minería de datos). España. Retrieved from http://masteres.ugr.es/moea/pages/tfm-1213/tfm_garciagonzalezfrancisco_1/
- Garcia D. 2016. Minería de datos aplicada a la enseñanza virtual: nuevas propuestas para la construcción de modelos y su integración en un entorno amigable para el usuario no experto.(Tesis de doctor), Universidad de Cantabria, Cantabria, España. Retrieved from

https://repositorio.unican.es/xmlui/bitstream/handle/10902/8551/Tesis_DGS.pdf?sequence=1&isAllowed=y

- Gutiérrez J. J. 2017. Propuesta de una metodología de extracción de conocimientos a partir de datos de las prestaciones del seguro integral del salud en la region Piura en el ano 2016. (Tesis de maestría), Universidad Católica Los Angeles de Chimbote, Piura, Perú. Retrieved from http://repositorio.uladech.edu.pe/bitstream/handle/123456789/2868/MINERIA_DE_DATOS_SERVICIO_DE_SALUD_GUTIRREZ_PACHERRES_JUAN_JESUS.pdf?sequence=1&isAllowed=y
- Hernández E. J. 2016. Aplicación de técnicas de análisis de datos y administración de Big Data ambientales. (Tesis de maestría), Universidad Nacional de Colombia, Medellin, Colombia. Retrieved from <http://bdigital.unal.edu.co/54512/1/1090175695.2016.pdf>
- Hidalgo B. 2018. Minería de datos en los Sistemas de gestión de Aprendizaje en la Educación Universitaria. Campus Virtuales, 7(2), 115–128. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=6681865>
- Holgado L. A. 2018. Detección de patrones de bajo rendimiento académico mediante técnicas de minería de datos de los estudiantes de la Universidad Nacional Amazónica de Madre de Dios 2018, (Tesis de maestría), Universidad Nacional del Altiplano, Puno, Perú. Retrieved from <http://repositorio.unap.edu.pe/handle/UNAP/9815>
- Marulanda E., Lopez M., Mejia M. 2017. Minería de datos en gestión del conocimiento de pymes de Colombia. Universidad Catolica Del Norte, 224–227. Retrieved from <http://www.redalyc.org/pdf/1942/194250865013.pdf>
- Melo A. D. 2018. Patrones para la estimación de consumo de medicamentos con minería de datos Redes Puno.(Tesis de maestría), Universidad Nacional del Altiplano, Puno, Perú. Retrieved from http://repositorio.unap.edu.pe/bitstream/handle/UNAP/9543/Alcides_Demetrio_Melo_Chura.pdf?sequence=1&isAllowed=y
- Mosquera R., Parra-osorio L., Castrillón O. D. 2016. Metodología para la Predicción del Grado de Riesgo Psicosocial en Docentes de Colegios Colombianos utilizando Técnicas de Minería de Datos. Información Tecnológica, 27(6), 259–272. <https://doi.org/10.4067/S0718-07642016000600026>
- Muñoz A. 2015. Modelos para la Mejora del Rendimiento Académico de Alumnos de la E.S.O. mediante Técnicas de Minería de Datos. (Tesis de doctor), Univesidad de Murcia. Murcia, España. Retrieved from https://idus.us.es/xmlui/bitstream/handle/11441/40436/Tesis_realizada_por_Daysi_K_García.pdf?sequence=1&isAllowed=y

-
- Ochoa L.L., Paredes K.R., Tejada J. E. 2017. Estudio Comparativo de Técnicas no Supervisadas de Minería de Datos para Segmentación de Alumnos. *Global Partnerships for Development and Engineering*, (July 2017), 19–21. <https://doi.org/10.18687/lacpei2017.1.1.115>
- Rojas Rodriguez, L. A. (2017). “Metamodelo para la integración de datos abiertos aplicado a Inteligencia de Negocios.” (Tesis de doctor), Universidad de Oviedo, Oviedo, España. Retrieved from <http://di002.edv.uniovi.es/~cueva/investigacion/tesis/Tesis-LuzAndrea.pdf>
- Ruiz E., Romero C. 2018. Resultados obtenidos en un proceso de minería de datos aplicado a una base de datos que contienen información bibliográfica referida a cuatro segmentos de la ciencia. *Journal of Information Systems and Technology Management – Jistem USP*, 15, 11. <https://doi.org/10.4301/S1807-1775201815003>
- Samaniego G., Marqués L., Gisbert M. 2014. Procedimiento para la definición de Patrones de Uso docente en un Sistema de gestión de Aprendizaje. *Revista Campus Virtuales*, III(2014), 88–95. Retrieved from https://www.researchgate.net/publication/273575522_Procedimiento_para_la_definicion_de_patrones_de_uso_docente_en_un_Sistema_de_Gestion_de_Aprendizaje
- Sinche A. 2016. El uso de las TIC's en el proceso de enseñanza - aprendizaje. *Educacion Universitaria*. Retrieved from https://www.researchgate.net/publication/314172999_El_uso_de_las_TICs_en_el_proceso_Ensenanza_-_Aprendizaje_Educacion_Universitaria
- Sulla J. A. 2015. Aplicación de técnicas supervisadas de minería de datos para determinar la predicción de deserción académica. (Segunda Especialidad), Universidad Católica de Santa María. Arequipa, Perú. Retrieved from <http://tesis.ucsm.edu.pe/repositorio/bitstream/handle/UCSM/3612/K7.0252.SE.pdf?sequence=1&isAllowed=y>
- Timarán S., Hernández I., Caicedo S., Hidalgo A., Alvarado J. C. 2016. Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional, 63–86. <https://doi.org/10.16925/9789587600490>
- Yamao E. 2018. Predicción del rendimiento académico mediante minería de datos en estudiantes del primer ciclo de la Escuela Profesional de Ingeniería de Computación y Sistemas, Universidad de San Martín de Porres, Lima - Perú. (Tesis de maestría) Universidad San Martín de Porres, Lima, Perú. Retrieved from http://www.repositorioacademico.usmp.edu.pe/bitstream/usmp/3555/3/yamao_e.pdf