

ISSN impreso: 1390-9754

PANORAMA

Revista Multidisciplinaria de la Universidad Católica de Cuenca



AÑO I • NÚMERO 1 • MARZO 2016



UNIVERSIDAD
CATÓLICA DE CUENCA
COMUNIDAD EDUCATIVA AL SERVICIO DEL PUEBLO

Fecha de recepción: 29 de octubre de 2015

Fecha de aceptación: 4 de enero de 2016

ALGORITMOS DE MINERÍA DE DATOS. UNA APLICACIÓN A EVALUACIÓN DOCENTE EN LA UNIVERSIDAD CATÓLICA DE CUENCA

DATA MINING ALGORITHMS. AN APPLICATION
TO TEACHING EVALUATION
IN THE CATHOLIC UNIVERSITY OF CUENCA

Segundo Leopoldo Pauta Ayabaca
Universidad Católica de Cuenca. Ecuador.

Correspondencia:

leopoldopautaa@gmail.com

spauta@ucacue.edu.ec

Ingeniero de Sistemas por la Universidad de Cuenca; Especialista en Docencia Universitaria UCACUE; Diplomado en Pedagogías Innovadoras UTPL; Diplomado en Metodologías de la Investigación por la Universidad Nacional Autónoma de México UNAM; tesista de Maestría en Administración de Bases de Datos de la Universidad Técnica de Ambato UTA; Autor de los libros: Lógica de Programación con Pseudocódigos y Flujogramas, 1999; Algoritmos Computacionales. Un enfoque Práctico, 2003, ISBN 978-9942-03-619-3; Algoritmos Computacionales. Un enfoque Práctico. 2da Edición., 2009, ISBN 978-9942-03-619-3; Libro de Trabajo de Algoritmos Computacionales, 2011, ISBN 978-9942-03-979-8; Libro de Trabajo de Algoritmos Computacionales. 2da. Edición, 2015, ISBN 978-9942-8553-3-6; Director de Carrera de Ingeniería de Sistemas 2011-2013, Director del Departamento de Gestión de Calidad; Profesor de la Universidad Católica de Cuenca en Algoritmos, Business Intelligence, Centros de Datos, Director y Profesor de la carrera de Análisis de Sistemas del Instituto de Tecnologías Sudamericano - Cuenca. Miembro del Consejo de Redes Académicas de Ingenierías de Sistemas (RECSIC) del Ecuador.

RESUMEN

El constante avance tecnológico ha permitido en todas las organizaciones recolectar y almacenar todo tipo de información, volúmenes crecientes de información quedan almacenadas, las cuales pueden brindar conocimiento. La minería de datos ofrece la posibilidad de llevar a cabo un proceso de análisis y descubrimiento de patrones que generen conocimiento y puedan brindar apoyo en la toma de decisiones, para mejorar y aumentar la productividad de las organizaciones. Estos análisis se fundamentan en algoritmos, que en base a una serie de cálculos y reglas heurísticas, permiten crear un modelo de minería a partir de los datos. En este contexto, se planteó una estructura investigativa con un enfoque cualicuantitativo, debido a la necesidad de análisis y medición de variables, así mismo se abordó las fases de la minería, tipos de algoritmos de minería de datos y su implementación con la herramienta WEKA, aplicando algoritmos de asociación en las bases de datos de evaluación docente de la Universidad Católica de Cuenca periodo 2014, primer semestre, cuyo objetivo fue encontrar patrones conductuales de comportamiento.

Palabras Claves: Patrones, Minería de Datos, Algoritmos, Asociación, WEKA, Reglas Heurísticas.

ABSTRACT

The ever-advancing technology has allowed in all organizations collect and store all kinds of information, increasing volumes of information are stored, which can provide knowledge. Data mining then offers the possibility of carrying out a process of analysis and discovery of patterns that generate knowledge and can provide support in decision-making to improve and increase the productivity of organizations. These analyzes are based on algorithms that

based on a series of calculations and heuristic rules let you create a mining model from the data. In this context an investigative structure with a quality-quantitative approach because of the need for analysis and measurement of variables was raised, also the phases of mining, types of data mining algorithms and their implementation with WEKA tool is addressed by applying algorithms association databases teacher evaluation of the Catholic University of Cuenca period 2014 first half, the goal was to find behavioral patterns of behavior.

Keywords: Patterns, Data Mining, Algorithms, Association, WEKA, Heuristics Rules.

1. INTRODUCCIÓN

La tecnología actualmente es parte de nuestra convivencia diaria, se encuentra inmersa en la industria, la banca, la educación en definitiva en el quehacer humano, en donde ha aportado innovaciones importantísimas, apoyando a la solución de problemas, facilitando el flujo de la información, obteniendo análisis instantáneos de información, datos actualizados en tiempo real, etc., lo que ha propiciado una gran cantidad de información difícil de asimilar. Science indica que para el 2007 existían 295 exabytes de bytes, en 2011 aproximadamente 600 exabytes es decir un trillón de bytes, inimaginable. Tomemos como referencia nuestros ordenadores saturados de información, en las organizaciones en las que laboramos abrumados de datos, información por doquier, y es allí donde debemos sacar partido, es allí donde entran en juego estas nuevas tecnologías informáticas. La minería de datos como un proceso que analiza y detecta cierto tipo de información oculta en grandes cantidades de datos, utilizando análisis matemáticos que permitan deducir patrones que están inmersos, no

visibles al ojo del ser humano, como por ejemplo, en un escenario de ventas, donde se analiza y determina los tipos de productos que se puedan vender juntos, su ubicación en cerchas de los supermercados, qué tipos de clientes compran un determinado producto, marcas preferidas, costos, etc. El objetivo es determinar secuencias de datos, que brinden información para la toma de decisiones.

El área de la educación actualmente tiene como herramienta fundamental a los sistemas de información en todas sus áreas. La academia ha sido una de las beneficiarias de las nuevas tecnologías, se utiliza actualmente en la gestión académica, es parte del proceso de enseñanza – aprendizaje, muchas de ellas basadas en la web, en los entornos virtuales, quedando registros en las bases de datos de esta información, que puede ser analizada para descubrir tendencias. A las organizaciones educativas les interesa conocer las razones de la deserción estudiantil por ejemplo, a los estudiantes cuáles son los mejores recursos y tareas que le permitan mejorar su aprendizaje, conocer la evaluación de los docentes (razón de este estudio), cómo éstos influyen en el proceso de aprendizaje, etc., todos bajo un mismo objetivo orientado a mejorar la calidad en la academia.

Estas interrogantes, pueden ser solventadas con las tecnologías de minería de datos.

La finalidad del KDD es:

- ☐ Procesar automáticamente grandes cantidades de datos “brutos”
- ☐ Identificar los patrones más significativos y relevantes.
- ☐ Presentarlos como conocimiento apropiado para satisfacer las metas del usuario.

DATO

INFORMACIÓN

CONOCIMIENTO

Tomado M, (2011) La Minería de Datos en educación (MDE)

Fig. 1 Finalidad KDD

(elaborado por el investigador)

La minería de datos es una fase dentro del Knowledge Discovery in DataBase – [KDD] – y se define como el “proceso de extracción de información previamente desconocida, válida y útil de grandes bases de datos y el uso de la información para tomar decisiones cruciales de negocios.” (Azoumana, 2013) .

II. DESARROLLO DE CONTENIDOS

A. Minería de Datos, sus Fases

Existen diferentes fases que se han tomado en cuenta para realizar minería de datos, sin embargo se ha referenciado algunos pasos específicos, que se aplicaron al caso de evaluación docente, partiendo desde la formulación de preguntas acerca de los datos hasta la creación de un modelo para responderlas, expuestas a continuación:

Definir el problema, que consiste en precisar el problema y considerar el uso de los datos, en el afán de proporcionar respuestas al problema.

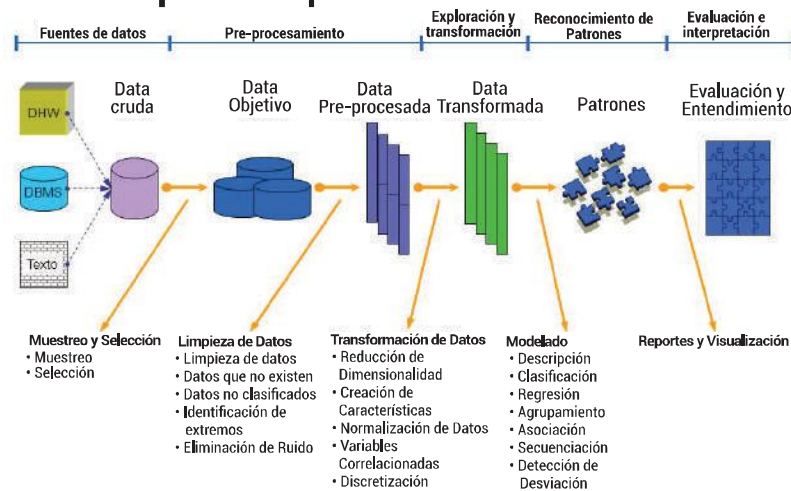
Recopilación de datos, en este proceso se recoge los datos disponibles de todas las fuentes posibles que intervienen en el factor de estudio, determinando los elementos claves que actúan en dicho proceso, es decir consolidar y limpiar los datos identificados en el problema. Debemos indicar que al referirnos a la consolidación, se refiere a identificar orígenes de datos que aporten de mejor manera a la solución.

Pre-procesado, al disponer con datos de distinto origen, se presenta un alta dimensionalidad y desbalanceo, por lo cual, es necesario realizar limpieza de datos, definición de variables (atributos) y re-balanceado de datos. En esta fase se aplicó técnicas de exploración, como son el cálculo de valores mínimos y máximos, cálculos de media,

desviación estándar, entre otros, lo que permitió determinar algunos datos que no son representativos, tal es el caso de la desviación estándar que apoyo determinando la estabilidad y la exactitud de los resultados.

Minería de datos, luego de agrupar y preparar la información, se define el método de minería a usarse, el cual estará sujeto al enfoque y el objetivo establecido en el estudio.

Mapa Conceptual - Minería de Datos



Interpretación de los resultados. En esta etapa se evalúa la información obtenida a través de la minería de datos, evaluando las relaciones entre factores, variables y elementos que interviene el caso de estudio propuesto. La figura 2 muestra el mapa conceptual del proceso de minería de datos.

Fig. 2 Mapa Conceptual – Minería de Datos

(elaborado por el investigador)

B. Algoritmos de Minería de Datos

“Un algoritmo de minería de datos es un conjunto de cálculos y reglas heurísticas, que permite crear un modelo de minería de datos a partir de los datos. Para crear un modelo, el algoritmo analiza primero los datos proporcionados, en busca de tipos específicos de patrones o tendencias. El algoritmo usa los resultados de este análisis para definir los parámetros óptimos para la creación del modelo de minería de datos.” (MSDN, 2015)

Uno de los aspectos más importantes a tomar en cuenta es determinar el tipo de algoritmo a utilizar, el cual dependerá de la visión del analista y los parámetros a determinar, a continuación describimos algunos tipos de algoritmos:

Algoritmos de clasificación.- Este tipo de algoritmo se basa en ciertos atributos del conjunto de datos, para predecir una o más

variables discretas. Algoritmos de regresión.- Este tipo de algoritmo se basa en ciertos atributos del conjunto de datos para predecir una o más variables continuas.

Algoritmos de segmentación.- Analiza los elementos que tienen propiedades similares para agruparlos.

Algoritmo de asociación.- Este algoritmo analiza diferentes atributos de un conjunto de datos en busca de correlaciones, generando reglas de asociación.

Algoritmo de análisis de secuencias.- El algoritmo busca ocurrencias frecuentes de datos.

C. WEKA. Software para Minería de Datos

WEKA es una herramienta de software basada en un conjunto de librerías JAVA, para la extracción de conocimientos desde

bases de datos. (José Hernández Orallo, 2006).

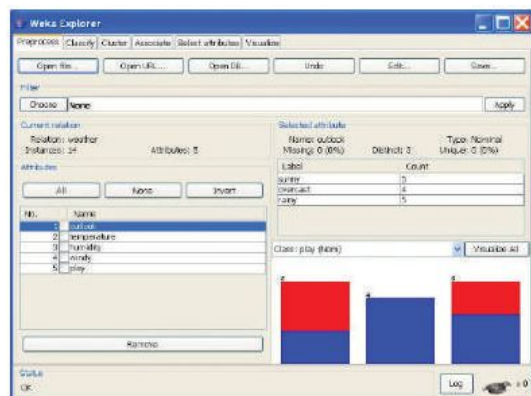


Fig. 3 Explorer de WEKA, con cargado de datos

(elaborado por el investigador)

Este tipo de herramienta se encuentra conformada por una serie de algoritmos, que permiten el análisis de datos y el modelado predictivo, apoyada de una interfaz gráfica que visualiza los resultados del preprocesamiento. WEKA tiene como base un conjunto de tareas esenciales para minería de datos, sobre todo en el preprocesamiento de datos como selección, visualización, clustering, clasificación y regresión. La herramienta para su utilización requiere de archivos planos (extensión .arff) determinado por registros con un número fijo de atributos, que al ser cargados son analizados inicialmente por WEKA determinado número de registros, atributos, su conformación, tipo de dato nominal o numérico, entre otros, que son presentados en un histograma con información de la distribución de datos, como se lo puede observar en la figura 3.

Una vez cargado el dataset es posible aprender un modelo, que puede ser parametrizado acorde a las necesidades en modelos de aprendizaje como: **Bayes**, fundamentado en el modelo de aprendizaje de su mismo nombre; **Funciones**, centradas en métodos matemáticos, regresiones, SVM entre otros; **Meta**, en donde se fusionan algunos métodos de aprendizaje; **Trees**, aplica la generación de árboles de decisión.

Definidos los modelos, WEKA comienza su aprendizaje en un determinado tiempo, dependiendo del dataset y los métodos aplicados. Al finalizar, la herramienta muestra información primero del dataset, luego de manera textual el modelo aprendido y finalmente información sobre la evaluación del modelo, que también es posible verlo de manera gráfica.

III. APLICACIÓN DE LA MINERÍA DE DATOS A LA EVALUACIÓN DOCENTE CON HERRAMIENTA WEKA

Acorde a las fases planteadas en el punto 2, ítem A, para el caso de estudio de evaluación docente se realizó lo siguiente:

1. Definición del Problema.

Actualmente las Instituciones Educativas Superiores, deberán dar cuenta constante a organismos rectores como es el CEAACES, de su constante gestión en el afán del aseguramiento de la calidad académica, debiendo evaluar, acreditar y asegurar la calidad de los programas y sus carreras. En este contexto, se ha presentado una serie de indicadores que las Universidades debe cumplir, de los cuales se ha tomado como referencia el indicador de *Eficiencia Terminal*, que aborda la tasa de graduación en función de cohortes e implícitamente niveles de deserción. En este contexto se analizarán los diferentes elementos que pueden influir en el cumplimiento de este indicador, tal es el caso de la evaluación docente y su influencia directa en el proceso de enseñanza, tratando de encontrar patrones en las preguntas realizadas a los estudiantes referente a las actividades de docencia.

2. Recopilación y Selección de Datos

Se realizó el análisis de la base de datos de evaluación docente de la Universidad Católica de Cuenca, correspondiente primer semestre del 2014.

La evaluación constó de un conjunto de preguntas agrupadas en función de: **Docencia** (Autoevaluación, Evaluación de Pares, Evaluación Directivos, Evaluación Estudiantil); **Investigación** (Autoevaluación, Evaluación de Pares, Evaluación Directivos); **Dirección y Gestión Académica** (Autoevaluación, Evaluación de Pares, Evaluación de Directivos, Evaluación Estudiantil). De las cuales se seleccionaron 29 preguntas que estaban directamente ligadas en el proceso de enseñanza.

Una vez determinado el conjunto de preguntas, se filtró los datos de tal manera que aminoren los valores incorrectos, los valores no válidos y desconocidos, permitiéndonos definir adecuadamente los datos a ser analizados.

3. Preprocesado

El objetivo de esta fase es aminorar la dimensionalidad y balanceo. Para poder cumplir con esta premisa, se determinó un grupo de preguntas que posiblemente incidan con fuerza en el indicador de eficiencia terminal, utilizando el método de elección de los mejores atributos del problema, obteniendo los siguientes resultados en la definición de variables:

Nro. de cédula, Sexo (m – f), Estado Civil (Casado – c –, Soltero – s –, Divorciado – d –), Rangos de Edad (30-40 años – X –, 40-50 años – Y –, más de 50 años – Z –), Título (cursando cuarto nivel, cuarto nivel, cursando PhD, PhD), Situación Laboral (Tiempo Completo – tc –, Medio Tiempo – mt –, Tiempo Parcial – tp –), Promedio docente (Excelente – E –, Muy Bueno – M –, Satisfactorio – S –, Regular – R –), y un grupo de 29 preguntas denominadas P, desde P1 – P29, tomadas del modelo de evaluación de la Universidad Católica de Cuenca, que se indica a continuación:

P1. ¿El syllabus contiene los resultados del aprendizaje a ser abordados en la asignatura?

P2. ¿El docente informa a los estudiantes los criterios y métodos de evaluación de los aprendizajes?

P3. ¿El docente diseña y utiliza material didáctico para facilitar el aprendizaje?

P4. ¿El docente como parte de los recursos didácticos, utiliza las TIC para facilitar el aprendizaje?

P5. ¿El docente elabora libros para contribuir a acrecentar el conocimiento y a mejorar el aprendizaje?

P6. ¿El docente demuestra que planifica y prepara bien las clases que imparte?

P7. ¿El docente demuestra que actualiza periódicamente la forma de impartir las clases?

P8. ¿El docente es puntual en la asistencia a clases?

P9. ¿El docente distribuye en forma adecuada el tiempo de trabajo en el aula, laboratorio o centro de prácticas?

P10. ¿Las actividades que desarrolla el docente en el aula, en el laboratorio o centro de prácticas, se ajusta a los objetivos y contenidos previstos?

P11. ¿El docente utiliza métodos y técnicas de enseñanza-aprendizaje adecuados?

P12. ¿Los materiales y demás recursos didácticos que utiliza el docente, facilitan su aprendizaje?

P13. ¿El docente demuestra tener formación y conocimientos sólidos de la asignatura que imparte?

P14. ¿Los contenidos e información que el docente imparte a los estudiantes como parte del syllabus es pertinente y actualizado?



P15. ¿El docente evidencia una buena coordinación entre las clases teóricas y prácticas que imparte?

P16. ¿El docente da respuestas satisfactorias a las inquietudes que formulan los estudiantes en clase?

P17. ¿El docente promueve la participación individual y grupal de los estudiantes en clase?

P18. ¿La bibliografía que proporciona el docente es actualizada, suficiente y pertinente a los contenidos que imparte en la asignatura?

P19. ¿Los contenidos que imparte el docente están relacionados con la realidad local y nacional?

P20. ¿El docente tiene excelente capacidad para comunicarse en forma verbal con sus estudiantes?

P21. ¿El docente reconoce los logros e iniciativas de los estudiantes?

P22. ¿El docente desarrolla con los estudiantes, actividades de refuerzo y transferencia de conocimientos?

P23. ¿El docente demuestra un elevado nivel de cumplimiento de la programación de la asignatura?

P24. ¿El docente, fuera del aula, orienta a los estudiantes a través de tutorías virtuales o presenciales, al desarrollo de sus tareas y demás dificultades de aprendizaje?

P25. ¿El docente, fuera del aula, realiza un acompañamiento programado a través de tutorías individuales o grupales a los estudiantes sobre el desarrollo de sus tareas y dificultades de aprendizaje?

P26. ¿El docente es ejemplo de honestidad, tolerancia y respeto?

P27. ¿El docente realiza visitas de campo y docencia en servicio?

P28. ¿El docente realiza la dirección, seguimiento y evaluación de prácticas académicas?

P29. ¿El docente realiza la dirección, seguimiento y evaluación de prácticas pre profesionales (servicios comunitarios) y/o pasantías profesionales?

A partir de la definición de las dimensionalidad y ajuste del balanceo preparó el archivo.arff, que se requiere para ejecutar en la herramienta WEKA, como se indica en la figura 4.

```

1 @relation EVALUACIONDOCENTE
2
3 @attribute Cedula String
4 @attribute Sexo {f, m}
5 @attribute EstadoCivil {c, d, s}
6 @attribute Edad integer
7 @attribute Titulo {te, cu, scu, ph, sph}
8 @attribute SituacionLaboral {tc, mt, tp}
9 @attribute P1 real
10 @attribute P2 real
11 @attribute P3 real
12 @attribute P4 real
13 .
14 .
15 .
16 @attribute P27 real
17 @attribute P28 real
18 @attribute P29 real
19 @attribute PromedioDocente real
20
21 $data
22 0101133270,f,c,60,ph,tc,100,100,94,78,20,88,100,100,98,100,100,89,10
23 0100957109,m,c,62,ph,tc,100,100,76,88,20,91,99,76,100,100,100,78,100
24 0102225505,m,c,46,cu,tc,85,99,81,89,82,75,76,85,87,98,100,94,76,99,100,9
25 1400116883,m,c,59,ph,tc,99,81,77,75,20,95,84,81,94,79,87,79,98,89,81,87,
26 0100191584,m,c,60,ph,tc,95,98,89,90,87,86,83,98,100,100,100,83,75,93

```

Fig. 4 Archivo .arff definido para WEKA

(elaborado por el investigador)

4. Minería de Datos

En esta fase, se aplicó una técnica de minería de datos, que permitió encontrar de un conjunto de modelos, uno que brinde conocimiento y que representa patrones de comportamiento encontrados en conjuntos de datos, plasmados en las variables del problema. En función de encontrar patrones de conocimiento y sugerencias de otros estudios, utilizamos Associate (denominado en WEKA) que usa reglas de asociación para el análisis de datos, en este caso puntual usaremos a priori. En el caso propuesto, se generó la relación entre el conjunto general de preguntas efectuadas a los estudiantes sobre los Docentes de la Universidad Católica en el primer semestre del 2014, aplicando factores como la edad,

sexo, estado civil y situación laboral de los docentes, esto permitió generar un modelo de asociación con una serie de conjuntos de elementos y de reglas que describen cómo estos elementos se agruparon dentro de los casos. Esas reglas que el algoritmo identificó, sirvió para encontrar preguntas excluyentes con bajo promedio, que aportarán en la toma de decisiones en el ámbito docente de la gestión.

5. Interpretación y Evaluación

Luego de realizar una serie de pruebas en la herramienta WEKA, se obtuvo el modelo final, lo que permitió comprobar que las conclusiones arrojadas sean válidas, suficientes y satisfactorias. Ya en esta última fase de interpretación y evaluación, utilizamos la opción Visualize que provee WEKA, donde pudimos estudiar el comportamiento de los datos mediante técnicas de visualización. Como se planteó en la primera fase, el objetivo es el de buscar incidencia en las evaluaciones docentes, es por eso que se agruparon las preguntas P25 y P28, P27 y P28, para tratar de encontrar patrones y se obtuvo lo siguiente:

Preguntas P25 y P28

P25. ¿El docente, fuera del aula, realiza un acompañamiento programado a través de tutorías individuales o grupales a los estudiantes sobre el desarrollo de sus tareas y dificultades de aprendizaje?

P28. ¿El docente realiza la dirección, seguimiento y evaluación de prácticas académicas?

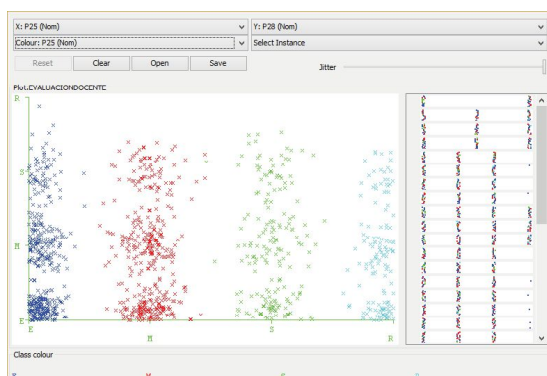


Fig. 5 Resultado de un nodo gráfico P25-P28 (K/S-K/R – PG) *(elaborado por el investigador)*

En la figura 5, se observa la combinación de las preguntas P25 y P28 en el entorno Visualize, donde debemos indicar que de entre todas las combinaciones posibles, el gráfico muestra algunas características significativas. Se observa una clara separación lineal entre K/S y K/E, de hecho existe una relación distante, que indica que la gestión del docente únicamente se realiza a nivel de aula, el docente no apoya significativamente a la gestión de prácticas preprofesionales y tutorías estudiantiles.

Preguntas P27 y P28

P27. ¿El docente realiza visitas de campo y docencia en servicio?

P28. ¿El docente realiza la dirección, seguimiento y evaluación de prácticas académicas?

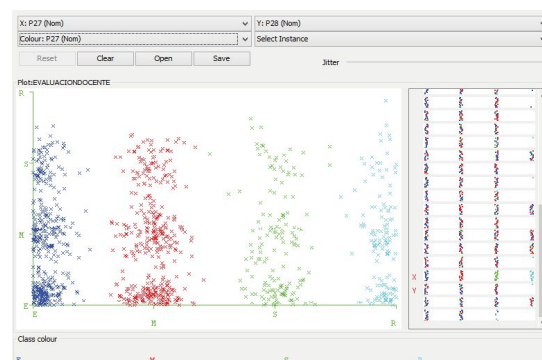
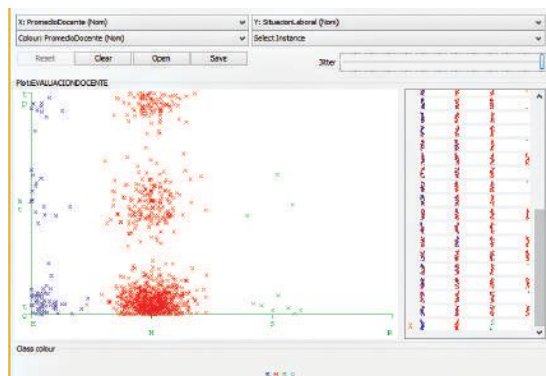


Fig. 6 Resultado de un nodo gráfico P27-P28 (K/S-K/R – PG) *(elaborado por el investigador)*

Al observar la combinación de las preguntas P27 y P28 en el entorno Visualize, el gráfico presenta características similares a la figura 6, corroborando el análisis de las preguntas P25 y P28, en donde no existe el apoyo docente extraclase.



En la figura 7, buscando patrones se combina la situación laboral (tc-mt-tp) con el promedio docente (Excelente – E –, Muy Bueno – M –, Satisfactorio – S –, Regular – R –), en donde se puede observar que el promedio docente de Muy Bueno (M) lo tienen la gran mayoría de docentes tc, y muy pocos docentes tiempo parcial tp, tienen una evaluación Excelente E.

Fig. 7 Resultado de un nodo gráfico P27-P28 (K/S-K/R – PG)

(elaborado por el investigador)

Revisión de reglas

Minimum support: 0.4 (440 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 12

Generated sets of large itemsets:

Size of set of large itemsets L(1): 32

Size of set of large itemsets L(2): 35

Size of set of large itemsets L(3): 5

Best rules found:

1. Edad=Y 473 ==> EstadoCivil=c 454 conf:(0.96)
2. Sexo=m SituacionLaboral=tc 483 ==> EstadoCivil=c 452 conf:(0.94)
3. Sexo=m EstadoCivil=c 560 ==> PromedioDocente=M 522 conf:(0.93)
4. Edad=X 484 ==> PromedioDocente=M 449 conf:(0.93)
5. Sexo=m SituacionLaboral=tc 483 ==> PromedioDocente=M 448 conf:(0.93)
6. Sexo=m 625 ==> PromedioDocente=M 579 conf:(0.93)
7. SituacionLaboral=tc PromedioDocente=M 711 ==> EstadoCivil=c 656 conf:(0.92)
8. EstadoCivil=c SituacionLaboral=tc 711 ==> PromedioDocente=M 656 conf:(0.92)
9. EstadoCivil=c 952 ==> PromedioDocente=M 877 conf:(0.92)
10. SituacionLaboral=tc 772 ==> EstadoCivil=c 711 conf:(0.92)

Fig. 8 Reglas que extrajo el algoritmo Apriori por defecto

(elaborado por el investigador)

Como se puede observar en la figura 8, se realizó un análisis de las reglas, con parámetros de cobertura en la parte izquierda y el nivel de confianza, así como también reglas que brinden mayor información, esta definición de parámetros permitió desechar algunas reglas que no aportaban conocimiento. Adicionalmente se agrupó reglas en función de conclusiones parecidas.

Planteamos entonces dos hallazgos interesantes:

Hallazgo 1.

R3. EstadoCivil=c P15=M 309 ==> PromedioDocente=M 300 conf:(0.97)

P15. ¿El docente evidencia una buena coordinación entre las clases teóricas y prácticas que imparte?

Los docentes con estado civil casado evidencian una mayor coordinación entre las clases teóricas y prácticas, a diferencia de los docentes solteros y divorciados.

Hallazgo 2.

R9. P29=M 405 ==> PromedioDocente=M 389 conf:(0.96)

P29. ¿El docente realiza la dirección, seguimiento y evaluación de prácticas pre profesionales (servicios comunitarios) y/o pasantías profesionales

Un alto porcentaje de docentes no realiza seguimiento a las prácticas pre profesionales, corroborando el resultado de la figura 5 y 6.

IV. CONCLUSIONES

Las IES deben implementar en sus centros de investigación métodos y técnicas de minería de datos, para obtener predicciones que le permitirá canalizar recomendaciones para la toma de decisiones oportunas.

Desarrollar algoritmos de minería de datos específicos, para problemas relacionados con la enseñanza y el aprendizaje.

La minería de datos está dirigida únicamente a personas con conocimientos en tecnologías y la gran mayoría las desconocen, por lo que deberían desarrollarse herramientas de minería más fáciles e intuitivas de utilizar, orientadas a personas no expertas.

Redefinir las bases de datos futuras, acorde a patrones encontrados al aplicar minería de datos.

CONFLICTO DE INTERESES

El autor declara no tener conflicto de intereses de ningún tipo.

IV. REFERENCIAS

Azoumana, K. (2013). Análisis de la deserción estudiantil en la Universidad Simon Bolívar, Facultad de Ingeniería de sistemas, con técnicas de minería de datos . Pensamiento Americano.

MSDN, M. (2015). Microsoft Developer Network. Obtenido de <https://msdn.microsoft.com/es-es/library/ms175595.aspx>

José Hernández Orallo, C. F. (2006). Introducción al WEKA The University of Waikato. València: The University of Waikato.

Alvaro Jimenes Galindo, H. A. Minería de Datos en la Educacion. Madrid: Universidad Carlos III de Madrid.

Lior Rokach, O. M. (2008). Data mining with decision trees: theory and applications. World Scientific. ISBN 978-981-277-171-1.

SENESCYT. (Mayo de 2011). elcomercio. Obtenido de http://www.elcomercio.ec/sociedad/desercion-mayor-publica_0_624537721.html

Maldonado, A. (2010). Estrategias para promover la retención estudiantil en un sistema de educación a distancia. Loja: Universidad Técnica Particular de Loja UTPL.

R. Alcover, J. B. Análisis del rendimiento académico en los estudios de informática de la Universidad Politécnica de Valencia aplicando técnicas de minería de datos. Valencia: Universidad Politécnica de Valencia.

Begueri Graciela, K. R. (2013). Determinación de factores resilientes en los alumnos universitarios. San Juan: UNSJ. XV Workshop de investigadores en ciencia de la computación.

Cristobal Romero, S. V. (2007). Educational data mining: A survey from 1995 to 2007. Expert System with Applications , 135-146.

Wnters, T. (2006). Collection and analysis of score matrices for outcomes-based assessment. Educational data mining:. California: University of California Riverside.