

MINERÍA DE DATOS DE LA EVALUACIÓN INTEGRAL DEL DESEMPEÑO ACADÉMICO DE LA UNIDAD DE NIVELACIÓN

***María Isabel Uvidia Fassler**

✉ muvidia@unach.edu.ec

****Andrés Santiago Cisneros Barahona**

✉ ascisneros@epoch.edu.ec

*****Javier Alonso Viñán Carrera**

✉ javinan@unach.edu.ec

Universidad Nacional de Chimborazo
Escuela Superior Politécnica de Chimborazo

RESUMEN

Las universidades ecuatorianas durante los últimos años han manejado un sinnúmero de procesos, siendo la evaluación integral al desempeño docente, importante y delicada. La Unidad de Nivelación y Admisión (UNAE) de la Escuela Superior Politécnica de Chimborazo (ESPOCH), mediante la aplicación del descubrimiento de conocimiento en base de datos (Knowledge Discovery in Databases - KDD), generó una propuesta metodológica acoplada a sus necesidades, denominada metodología KDD – ESPOCH, la misma que permitió generar una base de datos corporativa (Data Warehouse - DW) con información académica consistente. Para este análisis, se consideran los datos de las evaluaciones integrales; al aplicar algoritmos de clasificación de minería de datos (Data Mining - DM), árboles de decisión y redes bayesianas, como parte de la metodología, se generaron esquemas con porcentajes óptimos de confiabilidad, que permitieron conocer el desempeño de la actividad académica y participación de los docentes por género en las diferentes áreas del conocimiento y periodos académicos; establecer patrones y posibles tendencias de la evaluación integral; con la finalidad de identificar a las áreas del conocimiento de los cursos de nivelación que forman parte de la oferta académica de la institución, en las que se debe implementar estrategias de fortalecimiento académico y administrativo que permitan una incidencia positiva en el proceso enseñanza – aprendizaje.

PALABRAS CLAVES: Minería de datos, educación superior, evaluación docente, toma de decisiones.

ABSTRACT

The Ecuadorian universities during the last years have handled a number of processes, being the integral evaluation to the important and delicate educational performance. Facing this situation, the Standard Admissions Center (UNAE) of the Escuela Superior Politécnica de Chimborazo (ESPOCH), through the application of knowledge discovery in Databases (KDD), generated a methodological proposal adapted with the academic unit requirements, called the KDD - ESPOCH methodology, which allowed the generation of a data warehouse (DW) with consistent academic information. For this analysis, the data of the integral evaluations of the teachers of the UNAE are considered; (Data Mining - DM), decision trees and Bayesian networks, as part of the KDD methodology, schemas were generated with optimal percentages of reliability and precision, which allowed to know the performance of the academic activity and participation of teachers of this department by gender in the different areas of knowledge and by academic periods. Furthermore, to establish patterns and possible integral evaluation trends. Finally, the purpose was to identify the areas of knowledge in the UNAE courses with the academic offer of the university, in which academic and administrative strengthening strategies should be implemented that allow a positive impact on the teaching - learning process.

KEYWORDS: Data Mining, higher education, Integral evaluation to the teaching performance, making decisions.

1. INTRODUCCIÓN

La minería de datos o Data Mining (DM), es el “acto de extraer patrones o modelos a partir de los datos” Fayyad, Piatetsky - Shapiro y Smuth (1996), además es el proceso que consiste en extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos Hernández, Ferrari y Ramírez (2004).

Hoy, más que nunca, los métodos analíticos avanzados son el arma secreta de muchos negocios exitosos (Asencios, 2004), siendo la información veraz la que está al principio y al final de toda acción gerencial, por tal motivo, es importante, para la toma de decisiones contar con información relevante, sólida y oportuna, para cumplir objetivos y aplicar mejoras y acciones Fernández, Duarte, Hernández y Sánchez (2010).

Actualmente las instituciones de educación superior (IES), están generando una gran cantidad de datos que son almacenados en diferentes repositorios, lo que no ha permitido contar con información confiable en los procesos de las unidades académicas. Uno de los procesos académicos más significativos, es el de evaluación integral de desempeño del personal docente, como hace referencia el Reglamento codificado de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación

Superior, en su artículo 75.- “Ámbito y objeto de la evaluación.- La evaluación integral del desempeño se aplicará a todo el personal académico de las instituciones de educación superior públicas y particulares. La evaluación integral de desempeño abarca las actividades de docencia, investigación, y dirección o gestión académica”. Por tal razón, en este trabajo de investigación, se realizó la generación de conocimiento mediante la aplicación de algoritmos de clasificación de DM, árboles de decisión y redes bayesianas, en la Unidad de Nivelación y Admisión (UNAE) de la Escuela Superior Politécnica de Chimborazo (ESPOCH), permitiendo identificar la evaluación de los docentes en las diferentes áreas de conocimiento de los cursos de nivelación.

DESARROLLO

Descubrimiento de Conocimiento en Base de datos o Knowledge Discovery in Databases (KDD)

KDD es un análisis automático, exploratorio y de modelado de los repositorios de datos de gran tamaño (Fayyad et al., 1996), tiene la tarea de inferir información válida a partir de los datos. Sin embargo, la calidad de estos modelos está íntimamente relacionada con los datos almacenados, dependiendo de factores como la existencia de datos erróneos o el gran tamaño de la base de datos (Fernández et al., 2010). Maimon y Rokach (2010) definen KDD como un proceso organizado de identificación válida, novedosa, útil que genera patrones

comprensibles de los conjuntos de datos grandes y complejos que maneja una entidad de educación superior, además que este proceso KDD es iterativo e interactivo y lo define en nueve pasos.

DM es una de las fases más importantes dentro del proceso KDD, ya que es una experiencia práctica que muestra claramente la aptitud de las técnicas para resolver problemas empresariales. También, aporta en la resolución de problemas científicos, que impliquen el tratamiento de grandes cantidades de datos Aluja (2001)

Data Mining o Minería de Datos (DM)

DM implica un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias que son observadas al examinar grandes cantidades de información (Hernández et al., 2004). Además, mediante este proceso se genera un modelo útil para la predicción. Dicho modelo se construye teniendo como fundamento los datos que se encuentran en una base de datos, a los cuales se le ha aplicado algún algoritmo justamente con el fin de plantear un modelo. González (2005) En conclusión se podría decir que DM es un proceso que integra los datos de diferentes fuentes para, posteriormente, extraer un importante conocimiento, es decir, identificar información trascendente, valiosa y útil, a partir de lo cual las instituciones van a poder tomar alguna significativa decisión. Daza (2010)

Algoritmos de DM

Los algoritmos de clasificación de datos es el proceso por medio del cual se encuentra propiedades comunes entre un conjunto de objetos de una base de datos y se los cataloga en diferentes clases, de acuerdo al modelo de clasificación. Agrawal y Srikant (1994)

Existen varios algoritmos de clasificación como: rough sets, árboles de decisión, redes neuronales, redes bayesianas, algoritmos genéticos entre otros. Jiménez y Timaran (2015)

Árboles de Decisión

Son estructuras de forma de árbol que representan conjuntos de decisiones. Estas decisiones generan reglas para la clasificación de un conjunto de datos, las cuales explican el comportamiento de una variable con relación a otras, y pueden traducirse fácilmente en reglas de negocio. Son utilizados con finalidad predictiva y de clasificación. Por ejemplo, los árboles de decisión pueden emplearse para optimizar respuestas de campañas, identificar clientes potenciales o realizar evaluación de riesgos. Bernabeu (2010)

Redes bayesianas

Según Aluja (2001), consiste en representar todos los posibles sucesos en los que se está interesado mediante un grafo de probabilidades condicionales de transición entre sucesos. Puede codificarse a partir del conocimiento de un experto o puede

ser inferido a partir de los datos. Permite establecer relaciones causales y efectuar predicciones (Aluja, 2001).

Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior (Codificación)

TÍTULO IV EVALUACIÓN Y PERFECCIONAMIENTO DEL PERSONAL ACADÉMICO

CAPÍTULO I DE LA EVALUACIÓN INTEGRAL DE DESEMPEÑO DEL PERSONAL ACADÉMICO

Artículo 78.- Componentes y ponderación.-

Los componentes de la evaluación integral son:

1. Autoevaluación.- Es la evaluación que el personal académico realiza periódicamente sobre su trabajo y su desempeño académico.
- 2, Coevaluación.- Es la evaluación que realizan pares académicos y directivos de la institución de educación superior.
3. Heteroevaluación.- Es la evaluación que realizan los estudiantes sobre el proceso de aprendizaje impartido por el personal académico.

La ponderación de cada componente de evaluación será la siguiente:

1. Para las actividades de docencia: autoevaluación 10-20%; coevaluación de pares 20-30% y de directivos 20-30%; y heteroevaluación 30-40%.
2. Para las actividades de investigación: autoevaluación 10-20%; coevaluación de pares 40-50% y de directivos 30-40%.
3. Para las actividades de dirección o

gestión académica: autoevaluación 10-20%; coevaluación de pares 20-30% y directivos 30-40%; y heteroevaluación 10-20%.

Los resultados de la evaluación integral y de sus componentes serán públicos. En caso de que el personal académico combine actividades de docencia, investigación y gestión, la ponderación de la evaluación sobre cada una de las mismas será equivalente al número de horas de dedicación a cada una (CES, 2016).

2. MATERIALES Y MÉTODOS

La metodología utilizada fue KDD-ESPOCH, que acopló las nueve fases KDD a la Metodología HEFESTO v.2.0. (Metodología propia para la creación de un Data Warehouse o base de datos corporativa, DW). Las cuatro primeras fases KDD: aprender el dominio de la aplicación, selección y creación de un conjunto de datos sobre la que se realiza el descubrimiento, pre procesamiento y limpieza, y transformación de datos; se encuentran dentro del subproceso DW, que tiene como producto el DW creado con datos validados y consistentes; ya que las fuentes de datos para tomar decisiones eran diversas, fue necesario agruparlas en un solo repositorio o DW que permita aplicar algoritmos de DM. En las fases siguientes dentro del subproceso DM: elección de la tarea de DM apropiada, elección del algoritmo DM, empleando el algoritmo DM, se obtiene como producto

los patrones generados mediante la aplicación de los algoritmos DM, árboles de decisión y redes bayesianas en los datos de la UNAE - ESPOCH; finalmente en las fases: evaluación y uso del conocimiento

descubierto, se consiguió como producto el conocimiento que permite tomar decisiones a los usuarios académicos de la IES. La metodología KDD-ESPOCH que se usa, se muestra en la figura 1.

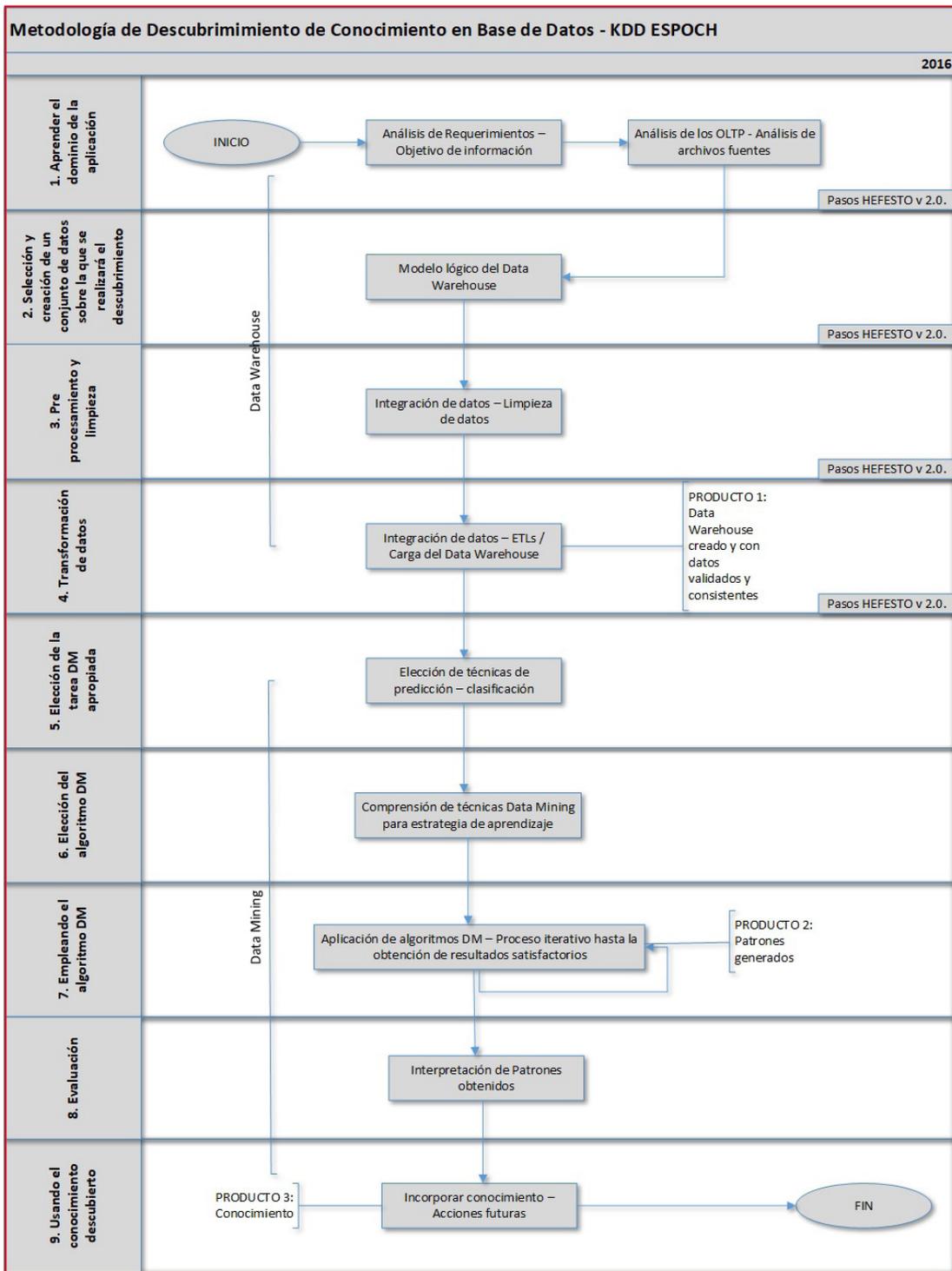


Figura 1: Metodología KDD implementado en la UNAE – ESPOCH.

Fuente: Propuesta de los autores

Elaborado: Por los autores

Diseño de la Investigación

El presente trabajo investigativo utiliza un diseño Cuasi Experimental, en el que se trabaja las evaluaciones integrales de desempeño del personal académico de la UNAE-ESPOCH. Estos archivos planos (Excel), mediante la aplicación de KDD-ESPOCH, se almacenaron en el DW para conseguir aplicar algoritmos DM y tomar decisiones académicas.

Métodos, Técnicas e Instrumentos

Para la implementación del DW se utiliza el gestor de base de datos PostgreSQL 9.3, que alojó un DW de tamaño personal con menos de 1GB. Para el análisis de los algoritmos DM, se aplicó WEKA Developer 3.7.13, permitiendo la obtención de patrones y conocimientos.

3. RESULTADOS

Mediante la aplicación de la metodología KDD-ESPOCH, se obtuvo un DW de tamaño personal, es decir de 1GB, que en su primer subproceso DW, consiguió almacenar los datos consistentes y confiables de la UNAE – ESPOCH. En el siguiente subproceso DM, se aplicaron algoritmos de clasificación: árboles de decisión y redes bayesianas.

Los datos del DW que se analizaron en este caso de investigación, fueron los datos de la evaluación integral del desempeño de los docentes de los cursos de nivelación de las diferentes áreas de conocimiento.

Los resultados de estas evaluaciones al desempeño docente de la UNAE evidencian la Heteroevaluación (Evaluación estudiante – docente, EST-DOC) con el 40% de la ponderación total, la Autoevaluación (Evaluación propia del docente) con el 20% de ponderación y la Coevaluación (Evaluación de los tutores) con el 40% de ponderación. Las ponderaciones permiten tener un puntaje total de 100 puntos; que fueron clasificados, como se muestra en la tabla 1, de acuerdo a los niveles de desempeño:

Nivel de desempeño	Rango	Descripción
Excelente	95% a 100%	El desempeño del docente es superior al esperado, cumple con todos los estándares y sus cualidades se observan en el dominio del saber que promueve, los conocimientos pedagógicos que emplea. Tienen nivel de liderazgo y confianza al promover la participación en el aula (INEVAL, 2016).
Favorable	85% a 94,99%	Su nivel de dominio disciplinar es adecuado, su desarrollo como docente y sus fundamentos pedagógicos le permiten planificar, ejecutar, evaluar el proceso de aprendizaje en un componente específico, y al mismo tiempo, afrontar situaciones adversas, resolver conflictos y fomentar la participación. Su nivel de liderazgo le ayuda a tomar decisiones a través de esquemas colaborativos que fomentan la participación en el aula (INEVAL, 2016).
Fundamental	75% a 84,99%	Posee conocimientos fundamentales y la noción de las habilidades pedagógicas con gestión del aprendizaje, mientras actúa como agente principal en el proceso educativo. Su nivel de liderazgo se ejerce en el aula con énfasis en el cumplimiento del currículo y la promoción de valores en el aula (INEVAL, 2016).
En formación	Inferior a 74,99%	Los conocimientos sobre su campo deben ampliarse y hacerse más profundos, a través de nuevas habilidades pedagógicas que le ayuden a planificar y ejecutar el proceso de aprendizaje, diseñar nuevas evaluaciones para un contexto específico. El fortalecimiento del liderazgo le ayudará a tomar decisiones, a través de esquemas colaborativos que le permitan afrontar situaciones adversas y fomentar la cultura de participación en el aula. (INEVAL, 2016)

Tabla 1: Nivel de desempeño docente y grado de desarrollo de un docente profesional.
 Fuente: Normativa para la evaluación del desempeño docente
 Elaborado: Instituto Nacional de Evaluación Educativa

Al aplicar el algoritmo árboles de decisión, se pudieron observar patrones, como se muestra en la figura 2. Estos pudieron identificar los conocimientos más relevantes de la evaluación de los periodos académicos 2015-1S, 2015-2S y 2016-1S.

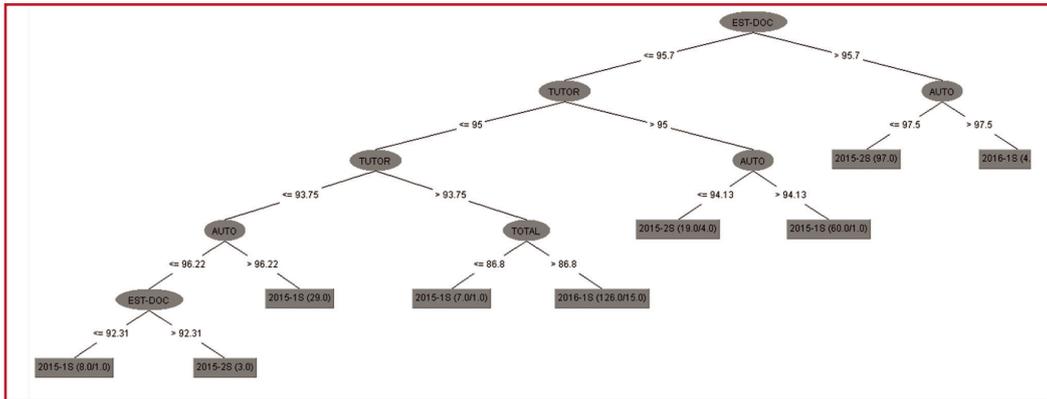


Figura 2: Árboles de decisión UNAE – ESPOCH.

Fuente: WEKA Developer 3.7.13

Elaborado: Por los autores

Los focos de conocimiento más importantes en el proceso de la evaluación del desempeño del docente, permite identificar que del total de 353 evaluaciones docentes, sólo 73 tienen una ponderación de excelente, 258 con ponderación favorable, 20 con ponderación fundamental y 2 con ponderación en formación; lo que implica un 21%, 73%, 5,5% y 0,5% respectivamente, y demuestra que la mayor cantidad de docentes se encuentran con ponderación favorable.

Además, en el periodo académico 2016-1S, de las evaluaciones EST-DOC (heteroevaluación) que tienen más del 95,7% (dato determinado por la concentración de

los puntajes de evaluación docente), sólo 4 docentes tienen en su autoevaluación más del 97,5%, y el resto tiene puntajes inferiores. También se puede observar, que mientras mayor sea la evaluación de los tutores (coevaluación), mayor es el puntaje de la autoevaluación del docente, y mientras menor sea la autoevaluación, los puntajes de EST-DOC (heteroevaluación) también serán menores. De esta forma se conocen las tendencias de las evaluaciones que determinan el puntaje global del desempeño de los docentes.

Al obtener patrones del algoritmo de redes bayesianas, como se muestra en la figura 3, se pudieron obtener varios conocimientos.

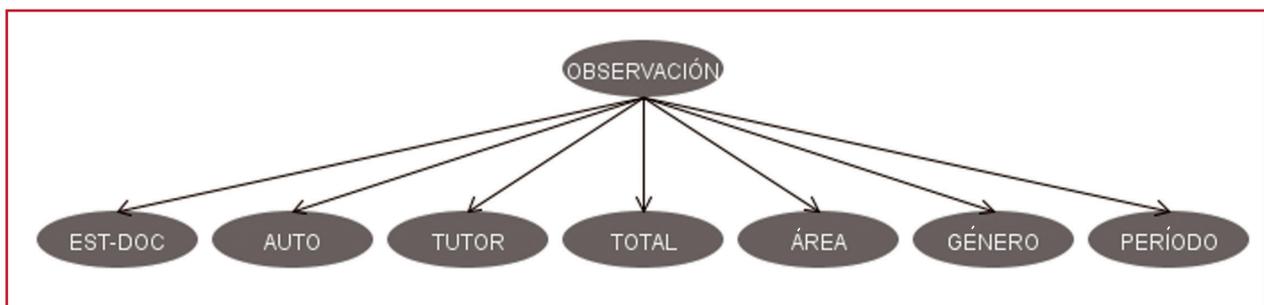


Figura 3: Redes bayesianas UNAE – ESPOCH.

Fuente: WEKA Developer 3.7.13

Elaborado: Por los autores

La participación de docentes del género femenino en el periodo 2015-1S fue del 44% frente al masculino del 56%, ésta tendencia se mantiene en los periodos académicos siguientes; en el 2015-2S, 41% de docentes son del género femenino y 59% del género masculino; y finalmente en el periodo 2016-1S, el 48% son docentes del género femenino y el 52% son docentes del género masculino. Lo que demuestra que existe mayor cantidad de docentes del género masculino.

El periodo académico con mejores evaluaciones EST-DOC (heteroevaluación) fue en el periodo académico 2015-2S, donde el 82% de docentes tuvieron puntajes de la evaluación al desempeño superior al 95%, es decir, de ponderación excelente. Además, la autoevaluación docente con ponderación excelente, fue del periodo 2015-1S con la representación del 80% de docentes, y la coevaluación más alta con nivel de desempeño excelente, lo consiguió el periodo 2016-1S con el 59% de docentes.

Respecto al desempeño global de los docentes de acuerdo al género, se identificó que los docentes del género femenino recibieron evaluaciones: 43% con ponderación excelente, 46% favorable, 31% fundamental y 5% en formación; mientras que los docentes del género masculino obtuvieron: 57% con ponderación excelente, 54% favorable, 69% fundamental y el 5% en formación. Lo que demuestra que, los docentes del

género masculino representan la mayoría en la planta docente de la UNAE, y que también son los que tienen mayor cantidad de evaluaciones con ponderación excelente, favorable y fundamental, siendo la ponderación fundamental la más grande, lo que les ubica en rangos de evaluación de 75% y 84,99%; mientras que el género femenino mantiene la mayor cantidad de evaluaciones en la ponderación favorable, con rangos entre 85% y 94,99%, notándose el mejor desempeño de las docentes mujeres.

Además, la concentración de ponderaciones por áreas se encuentra, Agricultura tiene el 68% de evaluaciones con ponderación favorable, 28% excelente, 2% fundamental y 2% en formación; Servicios tiene, el 48% en ponderación excelente y favorable, y 2% en ponderación fundamental y en formación; Salud tiene, el 72% en ponderación favorable, 18% excelente; Ingeniería tiene, el 69% en ponderación favorable, 20% excelente, 1% fundamental; Tronco común tiene, el 77% en ponderación favorable, 16% excelente, 6% fundamental, 1% en formación; Comercial tiene, el 82% en ponderación favorable y 16% excelente y 1% fundamental y en formación; Artes tiene, el 71% en ponderación favorable, 21% excelente, 4% fundamental, 4% en formación; Finalmente, de acuerdo a la concentración de los datos, la mayor cantidad de evaluaciones al desempeño docente se encuentra en la ponderación Favorable, es decir con rangos de 85% a 94,99%.

Las áreas mejor evaluadas: Servicios, Comercial y Salud; y con menores puntajes: Agricultura e Ingeniería.

4. DISCUSIÓN

En las IES es necesario implementar metodologías KDD, que permitan contar con datos confiables almacenados en un DW, de los cuales se pueda obtener información que mediante la aplicación de algoritmos DM apropiados en el área educativa generen conocimientos, que permitan conocer los escenarios de las diferentes unidades académicas, para plantear estrategias de mejora en el proceso de la evaluación al desempeño docente, como primera instancia a los docentes que laboran en la UNAE, y posteriormente, conocer la realidad de todas las carreras de la institución, con la finalidad de soportar la toma de decisiones.

5. CONCLUSIONES

- La implementación adecuada del proceso KDD-ESPOCH, permitió contar con un DW con información académica de los procesos, y mediante la aplicación de los algoritmos de clasificación de DM, árboles de decisión y redes bayesianas, se consiguió generar conocimiento del proceso de evaluación integral al desempeño del personal académico de la UNAE-ESPOCH.
- De acuerdo a los conocimientos generados, se pudo identificar que la cantidad de docentes del género masculino

es superior en todos los periodos académicos en comparación con el número de docentes del género femenino. Al hacer una clasificación por género, se genera conocimiento que evidencia un número mayor de evaluaciones en el rango de ponderación fundamental para los docentes del grupo masculino; en el caso de los docentes del género femenino, a pesar de presentarse como una minoría, se observa que se alcanza un mejor rendimiento en la evaluación al desempeño docente, ya que se evidencia un mayor número de evaluaciones con una escala de ponderación favorable. En las carreras de las áreas del conocimiento de Servicios, Comercial y Salud se aprecian las mejores ponderaciones en la evaluación integral al desempeño de los docentes.

- En las carreras del área de Agricultura e Ingeniería se identifica a las evaluaciones integrales docentes con menor rendimiento, lo que nos permite visualizar una de las razones por las que en estos campos existe la mayor cantidad de porcentaje de estudiantes desertores.
- Es necesario implementar y fortalecer las metodologías de enseñanza aprendizaje que requieren los docentes, con la finalidad de garantizar la adquisición de conocimientos significativos, el incremento del índice de retención de los estudiantes en las diferentes carreras y el adecuado desarrollo de los procesos académicos de la ESPOCH en las áreas del conocimiento específicas.

6. BIBLIOGRAFÍA

Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rules. In 20th International Conference on Very Large Data Bases (p. 487 a 499). Santiago de Chile.

Aluja, T. (2001). La Minería de Datos, entre la Estadística y la Inteligencia Artificial. *Qüestiió*, 25(3), 479–498.

Asencios, V. V. (2004). Data Mining y el descubrimiento del conocimiento. *Industrial Data*, 7(2), 83–86. Retrieved from <http://www.redalyc.org/resumen.oa?id=81670213>

Bernabeu, R. (2010). Hefesto, 146. Retrieved from <http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/hefesto-metodologia-propia-para-la-construccion-un-data-warehhttp://www.dataprix.com/data-warehousing-y-metodologia-hefesto/ii-hefesto-metodologia-propia-para-la-construccion-un-data-wa>

CES, C. de E. S. (2016). Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior. Ecuador.

Daza Vergaray, A. (2016). *Data Mining, Minería de Datos* (primera). Lima: Editorial Macro.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases, 17(3), 37. <https://doi.org/10.1609/aimag.v17i3.1230>

Fernández, T., Duarte, A., Hernández, R., & Sánchez, Á. (2010). GRASP aplicado al problema de la selección de instancias en KDD.

González, L. (2005). Una arquitectura para el análisis de información que integra procesamiento analítico en línea con minería de datos. Universidad de las Américas Puebla.

Hernández, J., Ferrari, C., & Ramírez, M. (2004). *Introducción a la Minería de Datos*. Madrid: Pearson Educación.

INEVAL. (2016). Normativa para la Evaluación del Desempeño Docente. Ecuador.

Jiménez, A., & Timarán, S. (2015). Caracterización de la deserción estudiantil en educación superior con minería de datos, 28(Diciembre), 447–463.

Maimon, O., & Rokach, L. (2010). Data Mining and Knowledge Discovery Handbook. In *Data Mining and Knowledge Discovery Handbook* (pp. 22–38). https://doi.org/10.1007/0-387-25465-x_2