

**DESCRIPCIÓN DE UN PROCESO KDD PARA CARACTERIZAR Y COMPARAR AL
PROFESORADO DE ESTADÍSTICA EN DOS NACIONES LATINOAMERICANAS.
AVANCE DE UN ESTUDIO COMPARATIVO**

Jesús Humberto Cuevas Acosta – Greivin Ramírez Arce

jesus.humberto.cuevas@outlook.com – gramirez@itcr.ac.cr

Instituto Tecnológico de Chihuahua II, México. Instituto Tecnológico de Costa Rica,
Costa Rica

Tema: IV.2 Formación y actualización del profesorado

Modalidad: CB

Nivel educativo: Medio (11 a 17 años)

Palabras clave: educación estadística, formación del profesorado, KDD, minería de datos

Resumen

En este artículo se presenta el avance de un estudio comparativo entre el profesorado que enseña estadística de séptimo a duodécimo grado en instituciones educativas en el estado de Chihuahua, México y la provincia de Cartago, Costa Rica. Ambas naciones han intentado promover en la última década el desarrollo de una alfabetización estadística en sus ciudadanos a través de la institución escolar, para lo cual han efectuado cambios en los planes y programas de estudio. Particularmente se hace énfasis en describir la metodología llamada “Extracción de Conocimiento en Bases de Datos” (Knowledge Discovery in Databases, KDD), y el diseño planteado para caracterizar al profesorado.

Antecedentes

Actualmente hay un consenso internacional respecto a la importancia de la estadística como parte integral en la formación de un ciudadano moderno. Consecuentemente se han formulado programas para desarrollar una alfabetización estadística de manera masiva. Habitualmente se reconoce a la institución escolar y al profesorado como ejes indispensables para alcanzar este propósito; a la escuela como contexto óptimo y al profesorado como mediador de aprendizajes.

En América Latina, las modificaciones curriculares efectuadas la última década reflejan el interés por promover el uso y entendimiento de la estadística en las actividades cotidianas. Destacan los casos de Costa Rica y México por la magnitud de las reformas realizadas y el interés por desarrollar programas de capacitación y actualización del profesorado que las enseña en educación primaria y secundaria. En el año 2012 el Ministerio de Educación Pública (MEP) de Costa Rica culminó con un proceso largo de modificaciones en el currículo escolar. En el caso particular de las matemáticas, los nuevos programas de estudio incluyeron tópicos de probabilidad y estadística en todos los ciclos de estudio, desde el primero hasta el diversificado, esto es, del primero al

duodécimo grado (MEP, 2012). En México, la Secretaría de Educación Pública (SEP), ha efectuado adecuaciones curriculares en educación básica. La Reforma Integral de la Educación Básica (RIEB) es la más reciente de todas, las adecuaciones iniciaron en 2004 y el proceso concluyó en 2011. Se establecieron estándares en cada disciplina, entre ellos algunos relacionados con el manejo de información y solución de problemas en los que intervienen fenómenos azarosos (DGDC, 2011).

Planteamiento del problema

Aun cuando en ambas naciones se ha invertido mucho tiempo, dinero y esfuerzo intelectual, las reformas no se han aceptado como se esperaba. Los profesores ven con recelo las adecuaciones recientes. En este trabajo se considera que esto se debe a varios factores, a saber, (1) son procesos de elaboración inconsultos en los que únicamente participan las élites académico-políticas; (2) No se conoce con certeza la realidad y problemática educativa que se vive en escuelas y aulas; (3) hay una propensión a seguir lineamientos y estándares educativos propuestos por organismos multilaterales de orientación eminentemente económica y financiera; (4) pareciera que se visualiza al profesorado como parte del problema educativo y no como medio de solución y mejoramiento de la educación.

En el caso particular de las nuevas exigencias disciplinares, los profesores observan la incorporación masiva de temas estadísticos y probabilísticos en los programas de estudio. También notan la implantación de mecanismos novedosos para evaluar aprendizajes en términos del cumplimiento de estándares y perciben un nuevo catálogo de recomendaciones para “alfabetizar correctamente” a sus estudiantes en estadística. Posiblemente, los profesores interpreten de forma distinta el término alfabetización estadística y dudan si tienen el conocimiento disciplinar requerido para hacer frente a la obligación de cumplir con los estándares establecidos. Quieren saber si habrá material de apoyo escrito, audiovisual o de otro tipo para auxiliarse en su práctica docente; esperan entender como deberán modificar sus estrategias didácticas y en qué medida habrán de usar tecnología informática; desean conocer los cambios en la forma de evaluar, medir y calificar el aprendizaje de sus alumnos; ansían saber si los van a formar o al menos capacitar en los temas nuevos y si recibirán asesoría pedagógica y disciplinar permanente por parte de alguna instancia gubernamental o institución de educación superior.

Probablemente, se descuidaron aspectos esenciales en la configuración y puesta en marcha de las reformas. Algunos ligados estrechamente a la formación profesional y pedagógica del profesorado, otros relacionados con cuestiones de tipo social y económico. Se considera que entre los más significativos se encuentran: (1) No determinaron con precisión el perfil profesional del cuerpo docente; (2) desconocen en qué medida los profesores cuentan con el dominio disciplinar exigido para cumplir con los nuevos programas de estudio; (3) Posiblemente no examinaron variables relacionadas con las condiciones laborales y sus posibles efectos en el desempeño laboral; y (4) faltaron más indagaciones que permitieran identificar, describir y comparar los métodos didácticos y de evaluación que conoce el maestro y contrastarlos con los recomendados a usar en los programas de reciente creación.

En suma, no se conoce al profesorado de forma integral, y eso es necesario si se requiere que ellos y sus estudiantes alcancen los estándares educativos establecidos en las nuevas reformas.

Propósito del estudio

En virtud de las características de esta problemática se consideró pertinente efectuar un estudio comparativo entre el profesorado que enseña estadística de séptimo a duodécimo grado en instituciones educativas de la Región Centro Sur del Estado de Chihuahua, México y la Provincia de Cartago, Costa Rica.

Para su realización, se partió de varias premisas globales claramente diferenciadas. La primera de ellas indica que por mejor intencionada, estructurada y elegante que pueda ser una reforma educativa, ésta no tendrá éxito si los profesores no adoptan las adecuaciones pedagógicas y de contenido para aplicarlas en el aula de clase (Fullan & Hargreaves, 2006). La segunda plantea la pertinencia de volver a valorar y apreciar el papel de los profesores como ente integrador en toda propuesta de reforma, ya que existe una gran experiencia acumulada por la extensa, fructífera y honesta trayectoria en el ámbito escolar de muchos de ellos. La tercera premisa sugiere que antes de iniciar un proceso de reforma y poner en práctica un programa de formación profesional, es necesario efectuar un diagnóstico del sistema educativo a la luz de las tendencias internacionales y las necesidades sociales y culturales del contexto en el que se inscribe.

Método

Debido a la naturaleza del propósito planteado, se consideró pertinente la utilización de un método mixto y utilizar un *Diseño Explicativo Secuencial* (DEXPLIS) porque en una primera fase se acopiarán y analizarán datos cuantitativos. Hernández, Fernández y Baptista (2010) hacen alusión a este diseño de investigación y señalan que una de las ventajas de este es que tiene etapas claramente diferenciadas. Los resultados obtenidos servirán para estructurar y afinar las actividades de la segunda fase que consistirá en la obtención de datos cualitativos. Posteriormente, los hallazgos de ambas fases se integrarán y analizarán globalmente en un reporte final.

El avance hasta hoy

Además de efectuar un análisis exhaustivo sobre tópicos relacionados con la educación estadística y su vínculo con el movimiento de mejora y cambio educativo, se examinaron con detenimiento diferentes tecnologías de utilidad para el estudio.

Se encontró que las peculiaridades de la problemática y el alcance del propósito planteado pueden zanjarse con el uso del proceso denominado *Descubrimiento de Conocimiento en Bases de Datos* (KDD por sus siglas en inglés). Seguir el proceso KDD permite encontrar perfiles y patrones en los datos que se recopilarán de cada uno de los participantes y, mediante la aplicación de algoritmos derivados del campo de la Inteligencia Artificial (IA), conocer el perfil del profesorado en las dimensiones: *social, disciplinar y pedagógica*. Alcanzar lo anterior proporcionará resultados útiles para que los administradores educativos conozcan a sus profesores, detecten tendencias y tengan información suficiente y de calidad para diseñar programas de actualización, capacitación y formación profesional y pedagógica congruentes con las necesidades educativas internacionales y especialmente de sus regiones. Es conveniente señalar que el proceso KDD consta de varias fases que es necesario seguir tantas veces como sea necesario para obtener modelos explicativos válidos (ver Figura 1).

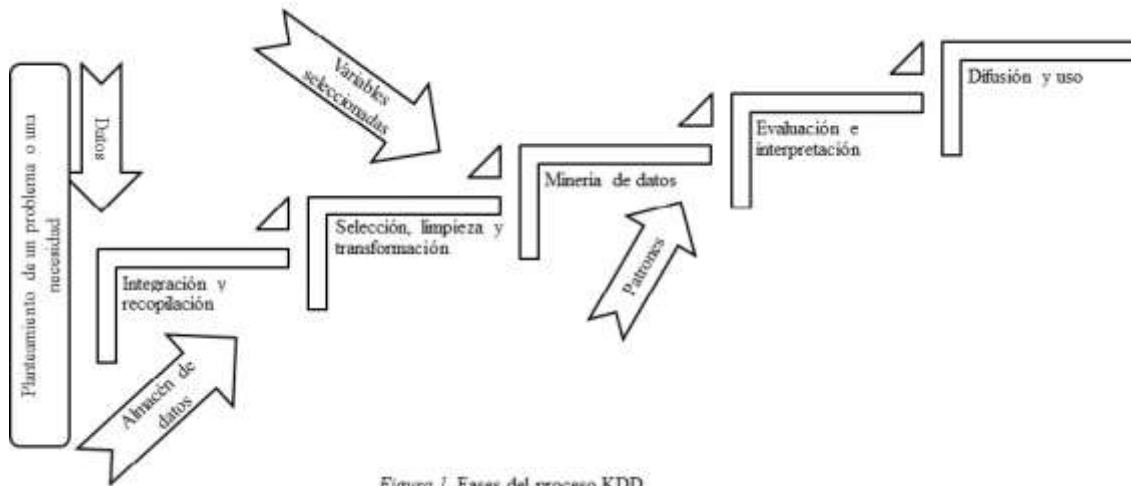


Figura 1. Fases del proceso KDD

Todas las fases son importantes, sin embargo, sobresale la de Minería de Datos (MD) por ser la que integra los algoritmos que se aplican al conjunto de datos con que se trabaja en cada proyecto. Es importante subrayar que popularmente se usan los conceptos de MD y KDD como sinónimos, lo cual es un error.

En este estudio, KDD se concibe como una metodología de trabajo integrada por fases diferenciadas pero relacionadas entre sí que permiten en su conjunto descubrir, identificar, sistematizar y representar numérica o gráficamente perfiles y patrones de comportamiento ocultos en un conjunto masivo de datos. Respecto al concepto de MD, se define como la fase de la metodología KDD en que se aplican métodos simbólicos y analíticos provenientes principalmente de la estadística, inteligencia artificial y sistemas de información, a un conjunto de datos sistematizados en fases previas.

Son muchos los campos en que se ha usado con éxito la MD, destacan la medicina, educación, mercadotecnia, seguridad financiera, industria y manufactura, meteorología, entre otros (Rodríguez & Díaz, 2009). En el caso de la educación, en los últimos años se ha utilizado para examinar problemas de diversa índole. En nivel superior, destacan los trabajos de Shyamala y Rajagopalan (2006), Goyal y Vohra (2012), Baradwaj y Pal (2012), Bhullar, Jaeng, y Kaur (2012). También se han realizado estudios para detectar patrones en el rendimiento de estudiantes, como los de Martínez, Acosta, Cutro, Uribe y Rambo (2010), Roldan (2012), Beguerí, Malberti, y Klenzi (2011). Los de Dapozo, Porcel, López, y Greiner (2008) y de Longoni, Porcel, López, y Dapozo (2010), tratan la misma problemática, pero se hace énfasis en las virtudes del uso de redes neuronales como alternativa a técnicas estadísticas tradicionales. Asimismo, debido a la

proliferación de la educación virtual, ha crecido el interés por analizar patrones de interacción entre los estudiantes, el trabajo de Valdiviezo (2011) es un ejemplo de la aplicación de MD en esa dirección.

Apuntes finales

En el caso de este estudio, se cuenta con guía de trabajo estructurada en tres grandes categorías: *social*, *pedagógica* y de *dominio disciplinar*. Cada una de ellas con interrogantes que demandan respuestas puntuales.

En lo social, las preguntas guía son: ¿En qué medida existe similitud entre los atributos sociales de los profesores de la Provincia de Cartago, Costa Rica y la Región Centro Sur del Estado de Chihuahua, México? ¿Cuál es el perfil profesional del profesorado que enseña estadística entre el séptimo y duodécimo grado? y, ¿Cuáles son las diferencias y similitudes en las condiciones laborales entre los profesores de ambas regiones?

En lo pedagógico, las preguntas son: ¿Qué diferencias y similitudes existen entre los métodos didácticos que están utilizando los profesores de ambas regiones para el tratamiento de tópicos estadísticos? ¿Cómo evalúan la calidad de los aprendizajes que logran sus aprendices? ¿En qué medida se han aceptado los nuevos programas de estudio por parte del gremio docente? y, ¿Cuál es la valoración real que ha efectuado dicho gremio respecto a los programas de capacitación diseñados para apoyarlos en la enseñanza de estos tópicos?

En el caso de la categoría relacionada con el dominio disciplinar, las interrogantes centrales son: ¿Cuál es el grado de dominio en las disciplinas de probabilidad y estadística por parte de profesores Cartagineses y Chihuahuenses? ¿Qué características comparten respecto de su dominio en estas disciplinas? ¿En qué medida dicho dominio es suficiente para cumplir con los estándares curriculares exigidos en las nuevas reformas educativas? ¿Cuáles son los textos y materiales de trabajo con mayor recomendación y uso en la formación académica de los futuros profesores de matemáticas y estadística?, y, ¿Cuáles son los textos sugeridos para la enseñanza en los nuevos programas?

Actualmente se trabaja en la elaboración y validación de varios instrumentos para recabar la información necesaria y dar respuesta a estas interrogantes. En el Anexo 1 se presentan algunas variables representativas a examinar en la dimensión *social* y la de *dominio disciplinar*. Como se dijo líneas atrás, en este trabajo se conceptualiza KDD como una metodología de trabajo en lugar de un proceso. Se espera que su seguimiento permita encontrar perfiles del profesorado de ambas regiones y ubicar patrones de comportamiento que puedan ser de ayuda para que los administradores educativos tengan información más completa que les sirva al momento de diseñar programas de actualización, capacitación y formación profesional y pedagógica congruentes con las necesidades educativas internacionales y especialmente de sus regiones.

Referencias bibliográficas

- Baradwaj, B., y Pal, S. (2012). Mining Educational Data to Analyze Students' Performance. *arXiv preprint arXiv:1201.3417*. <http://arxiv.org/abs/1201.3417> Consultado el 6/1/2013 Consultado el 24/04/2012
- Beguerí, G., Malberti, A., y Klenzi, R. (2011). Minería de datos en la determinación del perfil de un potencial alumno universitario desertor. *XIII Workshop de Investigadores en Ciencias de la Computación*. <http://sedici.unlp.edu.ar/handle/10915/20033> Consultado el 14/05/2012
- Bhullar, M. S., Iaeng, M., y Kaur, A. (2012). Use of Data Mining in Education Sector. *Proceedings of the World Congress on Engineering and Computer Science WCECS* http://www.iaeng.org/publication/WCECS2012/WCECS2012_pp513-516.pdf Consultado el 29/3/2013
- Dapozo, G. N., Porcel, E., López, M. V., y Greiner, C. L. (2008). Técnicas de clasificación aplicadas al estudio del rendimiento de ingresantes universitarios. *X Workshop de Investigadores en Ciencias de la Computación*. <http://sedici.unlp.edu.ar/handle/10915/20639> Consultado el 14/05/2012
- DGDC (2011). *Programas de estudio 2011. guía para el maestro*. http://basica.sep.gob.mx/dgdc/sitio/pdf/inicio/matlinea/2011/matematicas_sec.pdf Consultado el 16/01/2013
- Fullan, M. y Hargreaves, A. (2006). *La escuela que queremos. Los objetivos por los cuales vale la pena luchar*. buenos aires: Amorrortu editores.
- Goyal, M., y Vohra, R. (2012). Applications of Data Mining in Higher Education. *International Journal of Computer Science*, 9(2), 113-120
- Hernández, R., Fernández, C., y Baptista, P. (2010). *Metodología de la Investigación* (5a. ed.). Perú: McGraw-Hill.
- Longoni, M. G., Porcel, E., López, M. V., y Dapozo, G. N. (2010). Modelos de Redes Neuronales Perceptrón Multicapa y de Base Radial para la predicción del rendimiento académico de alumnos universitarios. *XVI Congreso Argentino de Ciencias de la Computación*. <http://sedici.unlp.edu.ar/handle/10915/19333> Consultado el 24/04/2013

- MEP (2012). *Reforma curricular en ética, estética y ciudadanía. Programas de estudio de matemática. I y II ciclo de educación primaria, III ciclo de educación general básica y diversificada.*
http://www.mep.go.cr/downloads/recursos tecnologicos/programa_matematicas.pdf
Consultado el 19/03/2013
- Martínez, D., Acosta, J., Cutro, L., Uribe, V., y Rambo, A. (2010). Data warehouse y data mining aplicados al estudio del rendimiento académico y de perfiles de alumnos. *XII Workshop de Investigadores en Ciencias de la Computación.*
<http://sedici.unlp.edu.ar/handle/10915/19461> Consultado el 14/1/ 2012
- Rodríguez, Y. y Díaz, A. (2009). Herramientas de Minería de Datos. *RCCI*, 3(34), 73-80.
- Roldán, M.(2012). Aplicación de una nueva metodología Adaptive Business Intelligence para un análisis taxonómico predictivo utilizado para la detección temprana de alumnos universitarios en riesgo de deserción. *UNLaR Ciencia*, (2).
<http://170.210.152.61/biblioteca/ojs-2.3.7/index.php/unlarciencia/article/view/205> Consultado el 14/05/2012
- Shyamala, K., y Rajagopalan, S. (2006). Data Mining Model for a Better Higer Educational System. *Information Technology Journal*, 5(3), 560–564.
- Valdiviezo, P. (2011). Aplicación de técnicas de aprendizaje automático para la identificación de patrones de interacción en una experiencia virtual de aprendizaje. *III Congreso CREAD ANDES y III Encuentro Virtual Educa Ecuador.* <http://repositorial.cuaed.unam.mx:8080/jspui/handle/123456789/2702>
Consultado el 30 de mayo de 2013

Anexo 1. Componentes a examinar en el estudio

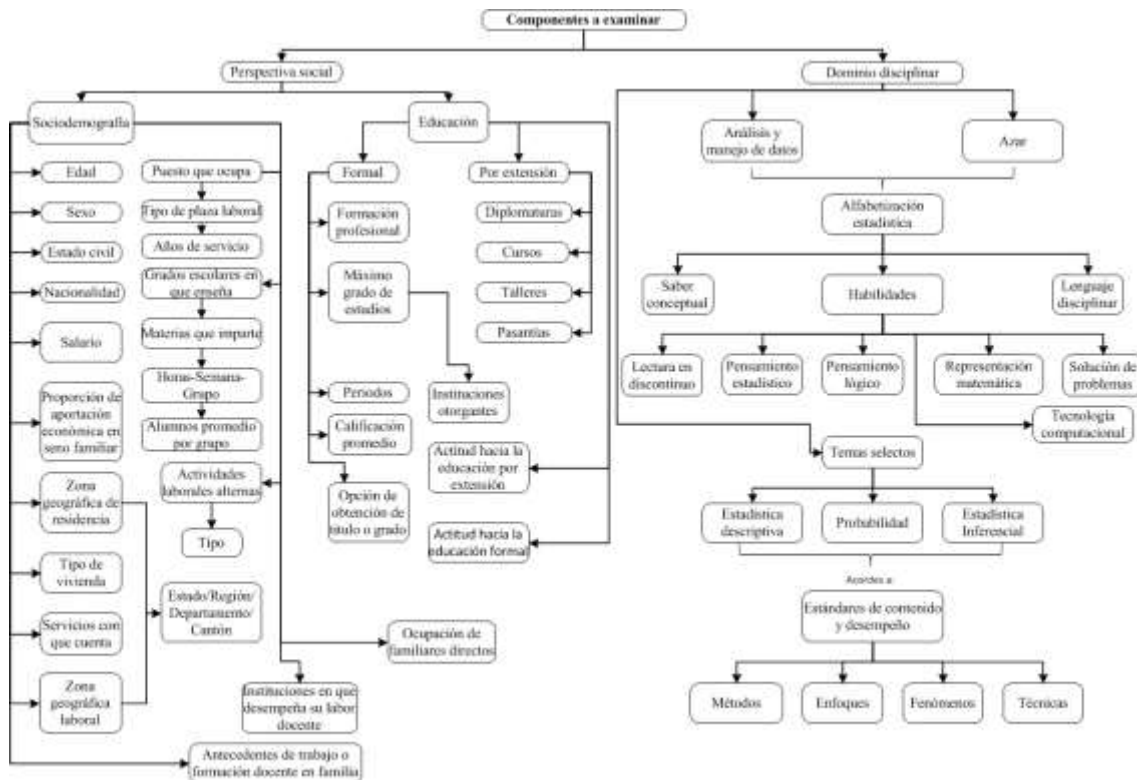


Figura 2. Componentes a examinar en el estudio