

Introducción

Uno de los objetivos del examen de Estado de calidad de la educación superior Saber Pro, según el Ministerio de Educación Nacional (MEN, 2009), a través del Decreto 3963 del 14 de octubre del 2009, es comprobar el grado de desarrollo de competencias de los estudiantes próximos a culminar los programas académicos de pregrado que ofrecen las instituciones de educación superior. El examen está compuesto por pruebas que evalúan competencias genéricas y específicas. De acuerdo con los lineamientos Saber Pro del Instituto Colombiano para la Evaluación de la Educación (ICFES, 2011a), todos los estudiantes deben presentar los módulos de competencias genéricas—sin importar el programa de formación que cursen—, que incluyen competencias de lectura crítica, comunicación escrita, razonamiento cuantitativo e inglés.

Según el ICFES (2012a), en la competencia de lectura crítica se evalúan los desempeños asociados a lectura, pensamiento crítico y entendimiento interpersonal; en comunicación escrita, la competencia para comunicar ideas por escrito referidas a un tema dado; en razonamiento cuantitativo, los desempeños relacionados con uso de lenguaje cuantitativo y solución de problemas; y en inglés, la competencia del estudiante para comunicarse efectivamente en inglés.

A pesar que en la prueba Saber Pro no se pretende que los estudiantes de todas las carreras profesionales desarrollen las competencias genéricas a un mismo nivel, ni aún las comunes a grupos de programas, sí es importante para el sector de la educación superior conocer cómo se asocian los factores correspondientes a las dimensiones sociodemográficas, económicas, académicas e institucionales del estudiante en relación con el nivel de desempeño en estas competencias según las pruebas Saber Pro.

Los estudios que se han realizado hasta el momento, según el ICFES (2010, 2011b y 2012b), Zapata (2011) y la Universidad Nacional de Colombia (2012), con respecto al análisis de los resultados de las pruebas Saber Pro, se basan en información procesada mediante un análisis estadístico, en el que fundamentalmente se consideran variables y relaciones primarias, sin tener en cuenta las verdaderas

interrelaciones, que por lo general están ocultas y que únicamente se pueden descubrir utilizando un tratamiento más complejo de los datos, lo cual es posible a través de la minería de datos. Mientras que la estadística plantea hipótesis que deben ser validadas a partir de los datos disponibles, la minería de datos descubre patrones no previstos con la estadística. Teniendo en cuenta a Febles y González (2002), los resultados que se obtienen con las técnicas de minería de datos no se pueden utilizar para confirmar o rechazar hipótesis, porque esto puede conducir a errores críticos. Su función es explorar datos, darles sentido, convertir en conocimiento un volumen de datos que poco o nada aportan a la descripción con el fin de interpretar fenómenos y adoptar decisiones de acuerdo con las necesidades.

La minería de datos en la educación no es un tema nuevo, su estudio y aplicación ha sido muy relevante en los últimos años; se pueden utilizar sus técnicas para explicar o predecir cualquier fenómeno dentro del campo educativo (Timarán, Calderón y Jiménez, 2013a y 2013b). Por ejemplo, utilizando las técnicas de minería de datos, se puede predecir, con un porcentaje muy alto de confiabilidad, la probabilidad de deserción de cualquier estudiante (Valero, 2009; Valero, Salvador y García, 2010). Las instituciones de educación pueden usar la minería de datos para hacer análisis comprensivos de las características de sus estudiantes, en métodos evaluativos, y develar así procesos exitosos o, por el contrario, detectar fraudes o inconsistencias (Orea, Vargas y Alonso, 2005).

El objetivo de este estudio fue descubrir patrones de desempeño académico en competencias genéricas de los estudiantes de programas profesionales en las pruebas Saber Pro 2011-2, a partir de los datos sociodemográficos, económicos, académicos e institucionales almacenados en las bases de datos del ICFES, utilizando técnicas de minería de datos. Se aplicó la metodología CRISP-DM y con los repositorios de datos —limpios y transformados para cada competencia— se descubrieron patrones asociados al buen o mal desempeño académico utilizando la técnica de clasificación basada en árboles de decisión. En los patrones descubiertos se destacan la acreditación institucional y la modalidad de estudio como dos atributos importantes asociados al desempeño académico de los estudiantes en las pruebas Saber Pro, en las cuatro competencias genéricas. El conocimiento descubierto se incorpora al existente y se puede integrar a los procesos de toma de decisiones del ICFES y de las instituciones gubernamentales y académicas que velan por la calidad de la educación superior.

La investigación que se reporta pretendió dar respuesta a la siguiente pregunta: ¿cuáles son los patrones sociodemográficos, económicos, académicos e

institucionales asociados al desempeño académico de los estudiantes en las competencias genéricas de los programas profesionales en las pruebas Saber Pro 2011-2?

En relación a la estructura, el libro está organizado por capítulos. En el capítulo I, se conceptualiza acerca de la calidad de la educación superior, particularmente sobre la acreditación institucional y la acreditación de programas, por ser uno de los patrones que muestran los resultados del estudio; igualmente se conceptualiza sobre el desempeño académico y las competencias genéricas en la formación de profesionales. En el capítulo II, se describen las diferentes etapas del proceso de descubrimiento de conocimiento en bases de datos y las fases de la metodología para proyectos de minería de datos CRISP-DM, utilizada en ese proyecto de investigación. Por su parte, en el capítulo III se hace un análisis exploratorio de los datos y se muestran algunas tendencias sobre el desempeño académico con relación a las competencias genéricas. El capítulo IV refiere el desarrollo del proyecto de investigación aplicando cada una de las fases de la metodología CRISP-DM, que orientó el desarrollo del proyecto. En el capítulo V se evalúan, interpretan y discuten los resultados obtenidos. Finalmente, en el capítulo VI, los investigadores presentan las conclusiones de la investigación con base en los resultados obtenidos y hacen recomendaciones tanto para el campo de la educación superior en lo relacionado con la política educativa de calidad de la misma, como para nuevas investigaciones, haciendo uso de la información recopilada en las bases de datos del ICFES y teniendo en cuenta las técnicas de la minería de datos.