

5 Resultados y discusión

Results and discussion

Resumen

Teniendo en cuenta que la investigación que se reporta se centró en la búsqueda de patrones de desempeño académico en las competencias genéricas de las pruebas Saber Pro 2011-2, en este capítulo se presentan los patrones más destacados en cada una de las competencias genéricas: lectura crítica, comunicación escrita, razonamiento cuantitativo e inglés. Dichos patrones se presentan como reglas para cada competencia y se listan en orden jerárquico según el soporte y la confianza de cada regla. En cuanto a la discusión, este capítulo presenta una reflexión sobre los patrones de mayor confianza, de los que se ubican por encima de la media y los que se sitúan por debajo de ella; este procedimiento se realizó para cada competencia genérica. Vale decir que entre los patrones descubiertos se destacan la acreditación institucional y la modalidad de estudio como dos atributos importantes asociados al desempeño académico.

Palabras clave: competencias genéricas, comunicación escrita, inglés, lectura crítica, razonamiento cuantitativo, patrones de desempeño.

Abstract

Considering that the reported research centered on the search for academic performance patterns in generic skills of the Saber Pro Exam 2011-2, this chapter shows the most outstanding patterns in each generic skill: critical reading, written communication, quantitative reasoning and English. Such patterns are presented as rules for each skill and listed in hierarchical order according to the support and reliability of each rule. Concerning the discussion, this chapter presents a reflection on the most reliable patterns of those above and those below the average; this procedure was performed for each generic skill. It should be noted that, among the patterns discovered, institutional accreditation and mode of study stand out as two important attributes associated with academic performance.

Keywords: generic skills, written communication, English, critical reading, quantitative reasoning, performance patterns.

¿Cómo citar este capítulo?/How to cite this chapter?

Timarán-Pereira, S. R., Hernández-Arteaga, I., Caicedo-Zambrano, S. J., Hidalgo-Troya, A. y Alvarado-Pérez, J. C. (2016). Resultados y discusión. En *Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional* (pp. 151-167). Bogotá: Ediciones Universidad Cooperativa de Colombia. DOI: <http://dx.doi.org/10.16925/9789587600490>



Resultados

En esta sección, por cada competencia se evalúan los modelos de clasificación basados en árboles de decisión y se interpretan los resultados obtenidos en la etapa de modelamiento.

Evaluación e interpretación de resultados para la competencia lectura crítica

Se analizaron los resultados obtenidos con el árbol de decisión construido con el conjunto de datos T97055A32LEC (véase figura 10), en el cual se almacenan los datos válidos sobre los factores sociodemográficos, económicos, académicos e institucionales de 97 055 estudiantes de programas profesionales que presentaron las pruebas Saber Pro 2011-2, en la competencia genérica de lectura crítica, en la que se escogió el atributo *mod_lectura_critica_desemp* como clase. Teniendo en cuenta lo anterior, se puede observar que este último clasifica correctamente a 62 781 instancias, que corresponde a un porcentaje de precisión del 64,7% y 34 274 instancias incorrectamente clasificadas, correspondientes a un porcentaje del 35,3% (véase figura 11).

Teniendo en cuenta la distribución de los valores del atributo clase *mod_lectura_critica_desemp* del repositorio T97055A32LEC que contiene 53 046 registros para el valor “sobre la media” y 44 009 registros para el valor “bajo la media”, y evaluando el modelo con la matriz de confusión de la figura 11, dicho atributo clasifica correctamente a 37 498 casos de estudiantes cuyos resultados en lectura crítica están sobre la media y a 25 283 casos que están bajo la media. Por otra parte, clasifica incorrectamente a 15 548 casos de estudiantes cuyos resultados en lectura crítica están sobre la media y 18 726 casos que están bajo la media. Esto significa que el modelo clasifica correctamente al 70,7% de los estudiantes que están sobre la media en lectura crítica y al 57,4% de las estudiantes que están bajo la media en esta competencia.

De acuerdo con las reglas de clasificación obtenidas con este modelo (véase figura 12), los patrones más representativos descubiertos, teniendo en cuenta el soporte y la confianza, son los siguientes:

- **Regla 2.** Si la IES no está acreditada y la metodología de formación es a distancia, el desempeño de los estudiantes en lectura crítica tiene mayor

probabilidad de estar bajo la media. El 12,48% de los 97 055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 69,11% de los 12 112 (12,48%) estudiantes se clasifica correctamente y el 27,52% de los 44 009 estudiantes que están bajo la media cumple este patrón.

- **Regla 5.** Si la IES no está acreditada, la metodología de formación es presencial, el carácter académico de la institución es de universidad, el número de IES en la zona OCAD a la que pertenece la institución está entre 60 y 70, el número de estudiantes de programas profesionales en la zona OCAD en la que se ofrece el programa está entre 10 000 y 20 000, los estudiantes pertenecen a programas no acreditados y además tienen personas a cargo, el desempeño de los estudiantes en lectura crítica tiene mayor probabilidad de estar bajo la media. El 4,85% de los 97 055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 62% de los 4710 (4,85%) estudiantes se clasifica correctamente y el 10,70% de los 44 009 estudiantes que están bajo la media cumplen este patrón.

- **Regla 4.** Si la IES no está acreditada, la metodología de formación es presencial, el carácter académico de la institución es de institución universitaria y los estudiantes están clasificados en el nivel 1 del Sisben, el desempeño de los estudiantes en lectura crítica tiene mayor probabilidad de estar bajo la media. El 2,07% de los 97 055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifican de esta manera. El 67,35% de los 2012 (2,07%) estudiantes se clasifican correctamente y el 4,57% de los 44 009 estudiantes que están bajo la media cumplen este patrón.

- **Regla 7.** Si la IES está acreditada, el desempeño de los estudiantes en lectura crítica tiene mayor probabilidad de estar sobre la media. El 16,80% de los 97 055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 78,87% de los 16 306 (16,80%) estudiantes se clasifica correctamente y el 30,74% de los 53 046 estudiantes que están sobre la media cumple este patrón.

- **Regla 9.** Si la IES no está acreditada, la metodología de formación es presencial, el carácter académico de la institución es de universidad, el número de IES en la zona a la que pertenece la institución está entre 60 y 70, y el número de estudiantes de programas profesionales en la zona OCAD en la que se ofrece el programa es mayor que 20 000, el desempeño de los estudiantes en lectura crítica tiene mayor probabilidad de estar sobre la media. El 13,93% de los 9

055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 62,80% de los 13 524 (13,93%) estudiantes se clasifica correctamente y el 25,49% de los 53 046 estudiantes que están sobre la media cumplen este patrón.

- **Regla 8.** Si la IES no está acreditada, la metodología de formación es presencial, el carácter académico de la institución es de universidad y el número de IES en la zona OCAD a la que pertenece la institución es mayor que 70, el desempeño de los estudiantes en lectura crítica tiene mayor probabilidad de estar sobre la media. El 8,37% de los 97 055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 69,43% de los 8126 (8,37%) estudiantes se clasifica correctamente y el 15,32% de los 53 046 estudiantes que están sobre la media cumple este patrón.

- **Regla 12.** Si la IES no está acreditada, la metodología de formación es presencial, el carácter académico de la institución es de universidad, el número de IES en la zona a la que pertenece la institución está entre 60 y 70, el número de estudiantes de programas profesionales en la zona OCAD en la que se ofrece el programa está entre 10 000 y 20 000, los estudiantes pertenecen a programas no acreditados, no tienen personas a cargo, no están clasificados en el Sisben y la institución es de tipo oficial, el desempeño de los estudiantes en lectura crítica tiene mayor probabilidad de estar sobre la media. El 2,80% de los 97 055 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 61,11% de los 2718 (2,80%) estudiantes se clasifica correctamente y el 5,12% de los 53,046 estudiantes que están sobre la media cumple este patrón.

Evaluación e interpretación de resultados para la competencia comunicación escrita

Se analizaron los resultados obtenidos con el árbol de decisión construido con el conjunto de datos T95337A32ESC (véase figura 13), en el cual se almacenan los datos válidos sobre los factores sociodemográficos, económicos, académicos e institucionales de 95 337 estudiantes de programas profesionales que presentaron las pruebas Saber Pro 2011-2 en la competencia genérica de comunicación escrita, en la que se escogió el atributo *mod_comunica_escrita_desemp* como clase. Teniendo en cuenta esto, se puede observar que dicho atributo clasifica correctamente a 56 669 instancias, que corresponden a un porcentaje de precisión del 59,4% y 38

668 instancias incorrectamente clasificadas, correspondiente a un porcentaje del 40,5% (véase figura 14).

Teniendo en cuenta la distribución de los valores del atributo clase *mod_comunica_escrita_desemp* del repositorio T95337A32ESC que contiene 46 986 registros para el valor “sobre la media” y 48 351 registros para el valor “bajo la media”, y evaluando el modelo con la matriz de confusión de la figura 14, se encuentra que este atributo clasifica correctamente a 25 435 casos de estudiantes cuyos resultados en comunicación escrita están sobre la media y a 31 234 casos que están bajo la media. Por otra parte, clasifica incorrectamente a 21 551 casos de estudiantes cuyos resultados en comunicación escrita están sobre la media y 17 117 casos que están bajo la media. Esto significa que el modelo clasifica correctamente al 54,13% de los estudiantes que están sobre la media en comunicación escrita y al 64,6% de las estudiantes que están bajo la media en esta competencia.

De acuerdo con las reglas de clasificación obtenidas con este modelo (ver figura 15), los patrones más representativos descubiertos, teniendo en cuenta el soporte y la confianza, son los siguientes:

- **Regla 3.** Si la IES no está acreditada y la metodología de formación es a distancia, el desempeño de los estudiantes en comunicación escrita tiene mayor probabilidad de estar bajo la media. El 12,36% de los 95 337 estudiantes evaluados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 62,42% de los 11 787 (12,36%) estudiantes se clasifica correctamente y el 24,38% de los 48 351 estudiantes que están bajo la media cumplen este patrón.
- **Regla 1.** Si la IES no está acreditada, la metodología de formación es presencial, los estudiantes son de género masculino y la zona OCAD en la que se ofrece el programa es Caribe, el desempeño de los estudiantes en comunicación escrita tiene mayor probabilidad de estar bajo la media. El 5,44% de los 95 337 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 64,92% de los 5188 (5,44%) estudiantes se clasifica correctamente y el 10,73% de los 48 351 estudiantes que están bajo la media cumple este patrón.
- **Regla 5.** Si la IES está acreditada y el número de estudiantes de programas profesionales en la zona OCAD en la que se ofrece el programa es mayor que 20 000, el desempeño de los estudiantes en comunicación escrita tiene mayor probabilidad de estar sobre la media. El 7,82% de los 95 337 estudiantes

analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 71,06% de los 7458 (7,82%) estudiantes se clasifica correctamente y el 15,87% de los 46 986 estudiantes que están sobre la media cumple este patrón.

- **Regla 6.** Si la IES está acreditada y el número de estudiantes de programas profesionales en la zona OCAD en la que se ofrece el programa está entre 10 000 y 20 000 y los estudiantes son de género femenino, el desempeño de los estudiantes en comunicación escrita tiene mayor probabilidad de estar sobre la media. El 5,04% de los 95 337 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 62,67% de los 4801 (5,04%) estudiantes se clasifica correctamente y el 10,22% de los 46 986 estudiantes que están sobre la media cumple este patrón.

- **Regla 8.** Si la IES no está acreditada, la metodología de formación es presencial, los estudiantes son de género femenino y el programa es acreditado, entonces, el desempeño de los estudiantes en comunicación escrita tiene mayor probabilidad de estar sobre la media. El 4,32% de los 95 337 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 60,51% de los 4115 (4,32%) estudiantes se clasifica correctamente y el 8,76% de los 46 986 estudiantes que están sobre la media cumple este patrón.

Evaluación e interpretación de resultados para la competencia razonamiento cuantitativo

Se analizaron los resultados obtenidos con el árbol de decisión construido con el conjunto de datos T97057A32CUA (véase figura 16), en el cual se almacenan los datos válidos sobre los factores sociodemográficos, económicos, académicos e institucionales de 97 057 estudiantes de programas profesionales que presentaron las pruebas Saber Pro 2011-2 en la competencia genérica de razonamiento cuantitativo, en la cual se escogió el atributo *mod_razona_cuantitativo_desemp* como clase. A partir de lo anterior, se puede observar que el atributo mencionado clasifica correctamente a 64 549 instancias, que corresponde a un porcentaje de precisión del 66,5% y 32 508 instancias incorrectamente clasificadas, correspondiente a un porcentaje del 33,5% (véase figura 17).

Teniendo en cuenta la distribución de los valores del atributo clase *mod_razona_cuantitativo_desemp* del repositorio T97057A32CUA que contiene 48 144 registros para el valor “sobre la media” y 48 913 registros para el valor “bajo la

media”, y evaluando el modelo con la matriz de confusión de la figura 17, este último clasifica correctamente a 29 093 casos de estudiantes cuyos resultados en razonamiento cuantitativo están sobre la media y a 35 456 casos que están bajo la media. Por otra parte, clasifica incorrectamente a 19 051 casos de estudiantes cuyos resultados en razonamiento cuantitativo están sobre la media y 13 457 casos que están bajo la media. Esto significa que el modelo clasifica correctamente al 60,43% de los estudiantes que están sobre la media en razonamiento cuantitativo y al 72,49% de los estudiantes que están bajo la media en esta competencia.

De acuerdo con las reglas de clasificación obtenidas con este modelo (véase 37 figura), los patrones más representativos descubiertos, teniendo en cuenta el soporte y la confianza, son los que se presentan a continuación:

- **Regla 1.** Si la IES no está acreditada, el área del grupo de referencia del programa es ciencias sociales y la metodología de formación es a distancia, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar bajo la media. El 10,23% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 73,88% de los 9927 (10,23%) estudiantes se clasifica correctamente y el 20,30% de los 48 913 estudiantes que están bajo la media cumple este patrón.
- **Regla 5.** Si la IES no está acreditada, el área del grupo de referencia del programa es ciencias sociales, la metodología de formación es presencial y el área del conocimiento al que pertenece el programa es ciencias de la educación, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar bajo la media. El 7,24% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 70,66% de los 7029 (7,24%) estudiantes se clasifican correctamente y el 14,37% de los 48 913 estudiantes que están bajo la media cumplen este patrón.
- **Regla 6.** Si la IES no está acreditada, el área del grupo de referencia del programa es ciencias sociales, la metodología de formación es presencial y el área del conocimiento al que pertenece el programa es ciencias sociales y humanas, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar bajo la media. El 6,13% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 68,35% de los 5953 (6,13%) estudiantes se clasifica correctamente y el 12,17% de los 48 913 estudiantes que están bajo la media cumple este patrón.

- **Regla 8.** Si la IES no está acreditada, el área del grupo de referencia del programa es ciencias humanas, la institución es de carácter privado y los ingresos familiares de los estudiantes es de tres salarios mínimos, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar bajo la media. El 4,29% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 66,54% de los 4166 (4,29%) estudiantes se clasifica correctamente y el 8,52% de los 48 913 estudiantes que están bajo la media cumple este patrón.
- **Regla 17.** Si la IES no está acreditada, el área del grupo de referencia del programa es ciencias naturales y técnicas, la metodología de formación es presencial, el programa no es acreditado y el género de los estudiantes es masculino, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar sobre la media. El 9,26% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 70,52% de los 8988 (9,26%) estudiantes se clasifica correctamente y el 18,67% de los 48 144 estudiantes que están sobre la media cumple este patrón.
- **Regla 12.** Si la IES está acreditada y el área del conocimiento al que pertenece el programa es ingeniería, arquitectura, urbanismo y afines, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar sobre la media. El 4,21% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 90,05% de los 4090 (4,21%) estudiantes se clasifica correctamente y el 8,50% de los 48 144 estudiantes que están sobre la media cumple este patrón.
- **Regla 16.** Si la IES está acreditada y el área del conocimiento al que pertenece el programa es economía, administración, contaduría y afines, el desempeño de los estudiantes en razonamiento cuantitativo tiene mayor probabilidad de estar sobre la media. El 3,91% de los 97 057 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 76,37% de los 3791 (3,91%) estudiantes se clasifica correctamente y el 7,87% de los 48 144 estudiantes que están sobre la media cumple este patrón.

Evaluación e interpretación de resultados para la competencia inglés

Se analizaron los resultados obtenidos con el árbol de decisión construido con el conjunto de datos T96946A32ING (véase figura 38), en el cual se almacenan los

datos válidos sobre los factores sociodemográficos, económicos, académicos e institucionales de 96 946 estudiantes de programas profesionales que presentaron las pruebas Saber Pro 2011-2 en la competencia genérica de inglés, en la cual se escogió el atributo *mod_ingles_desemp* como clase. A partir de esto se puede observar que dicho atributo clasifica correctamente a 69 600 instancias, que corresponden a un porcentaje de precisión del 71,8% y 27 346 instancias incorrectamente clasificadas, correspondientes a un porcentaje del 28,2% (véase figura 39).

Teniendo en cuenta la distribución de los valores del atributo clase *mod_ingles_desemp* del repositorio T96946A32ING que contiene 40 037 registros para el valor “sobre la media” y 56 909 registros para el valor “bajo la media”, y evaluando el modelo con la matriz de confusión de la figura 39, este clasifica correctamente a 20 764 casos de estudiantes cuyos resultados en inglés están sobre la media y a 48 836 casos que están bajo la media. Por otra parte, clasifica incorrectamente a 19 273 casos de estudiantes cuyos resultados en inglés están sobre la media y 8073 casos que están bajo la media. Esto significa que el modelo clasifica correctamente al 51,86% de los estudiantes que están sobre la media en inglés y al 85,81% de los estudiantes que están bajo la media en esta competencia.

De acuerdo con las reglas de clasificación obtenidas con este modelo (véase figura 40), los patrones más representativos descubiertos, teniendo en cuenta el soporte y la confianza, son los siguientes:

- **Regla 3.** Si la IES no está acreditada, la metodología de formación es presencial y los estudiantes están clasificados en el nivel 2 del Sisben, el desempeño de los estudiantes en inglés tiene mayor probabilidad de estar bajo la media. El 14,39% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 74,31% de los 13 951 (14,39%) estudiantes se clasifica correctamente y el 24,51% de los 56 909 estudiantes que están bajo la media cumple este patrón.
- **Regla 1.** Si la IES no está acreditada y la metodología de formación es a distancia, el desempeño de los estudiantes en inglés tiene mayor probabilidad de estar bajo la media. El 12,45% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 87,16% de los 12 074 (12 45%) estudiantes se clasifica correctamente y el 21,22% de los 56 909 estudiantes que están bajo la media cumple este patrón.
- **Regla 2.** Si la IES no está acreditada, la metodología de formación es presencial y los estudiantes están clasificados en el nivel 1 del Sisben, el desempeño de

los estudiantes en inglés tiene mayor probabilidad de estar bajo la media. El 10,69% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 79,66% de los 10 366 (10,69%) estudiantes se clasifica correctamente y el 18,22% de los 56 909 estudiantes que están bajo la media cumple este patrón.

- **Regla 4.** Si la IES no está acreditada, la metodología de formación es presencial, si los estudiantes no están clasificados en el Sisben y tienen además personas a cargo, el desempeño de los estudiantes en inglés tiene mayor probabilidad de estar bajo la media. El 8,90% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 67,79% de los 8627 (8,90%) estudiantes se clasifica correctamente y el 18,22% de los 56 909 estudiantes que están bajo la media cumple este patrón.

- **Regla 10.** Si la IES está acreditada y los estudiantes no tienen personas a cargo, el desempeño de los estudiantes en inglés tiene mayor probabilidad de estar sobre la media. El 1,0% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 81,72% de los 14 539 (15,0%) estudiantes se clasifica correctamente y el 36,31% de los 40 037 estudiantes que están sobre la media cumple este patrón.

- **Regla 12.** Si la IES no está acreditada, la metodología de formación es presencial, los estudiantes no están clasificados en el Sisben, no tienen personas a cargo, son de estrato 4 y el número de estudiantes de la zona en la cual se ofrece el programa es mayor que 20 000, el desempeño de los estudiantes en inglés tiene mayor probabilidad de estar sobre la media. El 2,30% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifica de esta manera. El 68,18% de los 2232 (2,30%) estudiantes se clasifica correctamente y el 5,57% de los 4 037 estudiantes que están sobre la media cumplen este patrón.

- **Regla 13.** Si la IES no está acreditada, la metodología de formación es presencial, los estudiantes no están clasificados en el Sisben, no tienen personas a cargo y son de estrato 5, el desempeño de los estudiantes en inglés tiene mayor probabilidad de estar sobre la media. El 2,05% de los 96 946 estudiantes analizados en esta competencia en la prueba Saber Pro 2011-2 se clasifican de esta manera. El 68,64% de los 1990 (2,05%) estudiantes se clasifica correctamente y el 4,97% de los 40 037 estudiantes que están sobre la media cumple este patrón.

Discusión

El objetivo de esta investigación 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 la técnica de clasificación basada en árboles de decisión. Dicho objetivo se logró en su totalidad.

Para efectos de la discusión de los resultados, se escogieron los patrones de mayor confianza, tanto los que se ubican por encima de la media, como aquellos que se sitúan por debajo de ella; este procedimiento se realizó para cada competencia genérica. En los patrones descubiertos, se destacan la acreditación institucional y la modalidad de estudio como dos atributos importantes asociados al desempeño académico.

En efecto, los modelos de clasificación basados en árboles de decisión, construidos para cada competencia genérica, muestran que el patrón de IES acreditadas se asocia con el buen desempeño académico de los estudiantes de programas profesionales en las cuatro competencias genéricas objeto de estudio. Este resultado está en línea con la fundamentación conceptual expresada en la política pública sobre calidad de la educación superior, particularmente sobre el tema de acreditación institucional, pues de hecho el cumplimiento de los doce factores de calidad señalados por el CNA (2013) da cuenta de las condiciones de favorabilidad de todos los procesos en el interior de las IES, que permiten un mejor desempeño académico de los estudiantes.

En la investigación sobre compromiso estudiantil y desempeño académico universitario, realizada en siete universidades colombianas con acreditación institucional de alta calidad, llevada a cabo por Pineda y Pedraza (2011), desde la perspectiva de los estudiantes, señala que estas universidades se preocupan por proponer actividades y tareas que implican la aplicación de conceptos y teorías en problemas prácticos; le asignan importancia al trabajo en equipo, utilización de medios electrónicos con fines académicos y a procesos de realimentación del aprendizaje de los estudiantes; fomentan espacios de socialización que fortalecen la calidad de las relaciones estudiante y comunidad educativa. Estas estrategias favorecen el desempeño académico de los estudiantes en el desarrollo de las competencias genéricas.

En el mismo orden de ideas, si la IES no está acreditada, se encontró que este hecho junto con la modalidad de estudio a distancia se constituyen en patrón que se asocia con el bajo desempeño académico de los estudiantes de carreras profesionales en las competencias genéricas de las pruebas Saber Pro 2011-2.

La modalidad del programa, según resultados de la investigación de Parra et al. (2012), realizada con estudiantes de primer semestre de pregrado de la Facultad de Ingeniería de la Universidad de Antioquia, cohorte 2012-2, hace referencia a la metodología mediante la cual se oferta y desarrolla el programa, que puede ser: presencial, semipresencial, a distancia y virtual. En dicho estudio se encontró que la modalidad del programa se asocia de manera importante al desempeño académico. Por ejemplo, en la condición de insuficiente se ubica el 17% de los matriculados en la modalidad presencial y el 50% de los de la modalidad virtual; cifras que indican que los estudiantes de programas presenciales tienen mejor desempeño que aquellos que están en la modalidad a distancia.

Por su parte, Artunduaga (2008), en relación con la modalidad de formación, considera que es factor clave del desempeño académico la forma como se orienta al estudiante en el aprendizaje. Al respecto, Vásquez y Rodríguez (2007) evidencian el papel del estudiante de modalidad a distancia y aseguran que es él el responsable de su éxito educativo, que le corresponde autodirigirse y autoorganizarse, con el fin de alcanzar los objetivos propuestos en el plan de estudio. Dichos autores argumentan que si no se tienen desarrolladas estas capacidades de independencia, tendería al bajo desempeño, pues su característica principal está dada por la separación entre estudiante y docente, que puede conllevar de manera paradójica, las ventajas y desventajas de este modelo y potenciar igualmente efectos desfavorables si los programas no están bien estructurados y diseñados.

En el mismo sentido, Rubio (2009) considera que la modalidad a distancia, a diferencia de la presencial, tiene sus propias características: autonomía en el aprendizaje, utilización de tecnología, apoyo-tutorial, comunicación masiva, entre otros, que influyen en el desempeño académico. Sarramona (2002) señala que la adaptación del material para la educación a distancia con todos los principios pedagógicos que esta requiere constituye una tarea difícil, dada la falta de especialistas en el tema. A menudo se confunde un recurso para la autoformación con un simple material informativo, lo cual posteriormente se convierte en factor negativo en el desempeño académico del estudiante y fuente de fracaso de los programas. El autor afirma que el bajo desempeño académico en la modalidad a distancia puede provenir de varias fuentes: capacidad intelectual insuficiente,

débil compromiso académico y otros factores provenientes de las dimensiones sociodemográfica y económica, que caracterizan al estudiante de modalidad a distancia. En los sociodemográficos destaca la edad, el grupo familiar, las personas a cargo, el nivel cultural, el estado civil, la ubicación geográfica, entre otros; en los factores económicos subraya el estrato social, los ingresos económicos individuales y familiares. Por su parte, Moncada y Rubio (2011) exponen, además, que estos estudiantes tienen características particulares que dependen de la dispersión geográfica, edad, estado civil, motivación, intereses profesionales, falta de tiempo, aislamiento, falta de recursos y manejo de la tecnología; la mayoría posee obligaciones laborales y familiares.

Particularmente, para cada competencia se encuentra que la acreditación institucional y la modalidad de estudio se asocian al buen o bajo desempeño académico de los estudiantes que presentaron las pruebas Saber Pro 2011-2.

Analizando el modelo de la figura 10, en la competencia de lectura crítica, el buen desempeño se asocia al hecho que la IES esté acreditada; el hecho de que la IES no lo esté, junto con la modalidad presencial, constituye un patrón de buen desempeño. Siguiendo el modelo jerárquicamente desde el nodo de modalidad presencial, se encuentra que el buen desempeño académico en esta competencia está asociado, además, con altos ingresos familiares (dimensión económica), número alto y medio de IES, número de estudiantes en la zona y con programas acreditados (dimensión institucional).

Por su parte, el bajo desempeño en lectura crítica tiene relación con IES no acreditadas y la modalidad a distancia. En el caso de la modalidad presencial, el bajo desempeño se asocia a la dimensión económica, bajo nivel de Sisben y con la dimensión institucional, las instituciones tecnológicas o técnicas profesionales. En este contexto, Garbanzo (2007) indica que factores como la pobreza y la falta de apoyo social están relacionados con el desempeño académico; no obstante, advierte que no hay correspondencia estricta entre las desigualdades sociales y educativas; indica que existen otros factores como la familia, el sistema educativo o la institución, que pueden incidir en tal desigualdad. Seibold (2000) señala que si bien el contexto socioeconómico afecta el nivel de calidad educativa, de ningún modo lo determina; así mismo, reconoce que se presenta asociación entre estas variables, pero no configura una relación causa-efecto.

Analizando el modelo de la figura 13, en la competencia de comunicación escrita, el buen desempeño se asocia al hecho de que la IES esté acreditada y que haya un alto número de estudiantes en la zona geográfica de la IES. En un orden

inferior de jerarquía, se encuentra que también tienen buen desempeño las mujeres y los hombres que estudien carreras de ciencias humanas. En el caso de que la IES no esté acreditada y la modalidad sea presencial, también es un patrón de buen desempeño el hecho de que el programa esté acreditado.

Por su parte, el bajo desempeño en comunicación escrita tiene relación con IES no acreditadas y con la modalidad a distancia. En el caso de la modalidad presencial, el bajo desempeño se asocia con el género masculino de las zonas geográficas Caribe, Centro sur y Pacífico.

En este contexto, están los hallazgos de Gómez y Soares (2013), quienes en su análisis sobre la diferencia de género en relación con el desempeño académico expresan que tales diferencias no son significativas y que no se puede afirmar de modo definitivo que exista una relación directa entre el rendimiento académico y el sexo; sin embargo, según Garbanzo (2007), las mujeres sí tienen un ligero mejor desempeño académico que los hombres. En el mismo sentido, Montero y Villalobos (2004) señalan que existe una relación significativa entre el sexo y el desempeño académico que está a favor de las mujeres.

Analizando el modelo de la figura 16, en la competencia de razonamiento cuantitativo, el buen desempeño se asocia al hecho de que la IES esté acreditada y que el área del grupo de referencia sea ciencias naturales y técnicas. En un orden inferior de jerarquía, se encuentra que también tienen buen desempeño los estudiantes de IES no acreditadas, de modalidad presencial y de ciencias naturales y técnicas. Por su parte, el bajo desempeño en la competencia de razonamiento cuantitativo tiene relación con IES no acreditadas y el área del grupo de referencia de ciencias sociales y humanas.

En efecto, los estudiantes de las áreas de ciencias sociales y humanas, y ciencias de la educación presentan bajo desempeño en esta competencia; realidad que es reconocida por las IES, a tal punto que como alternativa de solución vienen implementando cursos adicionales a sus planes de estudio con el fin de proporcionar una formación básica en matemáticas, en programas que al margen de las pruebas Saber Pro habría sido impensable su implementación. No obstante, lo ideal sería que aquellos cursos fueran constitutivos de los currículos de todos los programas de formación profesional, en la perspectiva de recuperar su aporte a la formación y estructuración del pensamiento matemático, de modo que, como señalan Davidson y McKinney (2001), se desarrolle la competencia para la interpretación e inferencia a partir de modelos; la representación de la información matemática de diversas formas; la utilización de métodos aritméticos, algebraicos,

geométricos y estadísticos para solucionar problemas; la estimación y chequeo de las respuestas a problemas matemáticos con miras a establecer su razonabilidad, identificar alternativas y seleccionar los mejores resultados; y el reconocimiento de los límites de los métodos matemáticos y estadísticos, para contribuir al fomento y desarrollo de la competencia de razonamiento cuantitativo.

Se entiende que es en este sentido como el Ministerio de Educación Nacional (s. f), concibe la alfabetización cuantitativa como el dominio del lenguaje y los métodos matemáticos indispensables para lograr la comprensión de problemas y discusiones cotidianas contemporáneas, sin lo cual las personas se encontrarían parcialmente incomunicadas. En la misma línea, haciendo referencia al pensamiento matemático como una competencia genérica emergente, los autores Toro y Villaveces (2008) plantean que “podemos pensar que hay un sustrato mínimo de competencias matemáticas, sin el cual un profesional puede quedar parcialmente incomunicado del mundo moderno que lo rodea” (p. 1); afirmación que muestra que la competencia del razonamiento cuantitativo es indispensable y, por lo mismo, inaplazable para todo profesional.

A pesar del reconocimiento de la importancia del razonamiento cuantitativo en la formación de todo profesional, se encuentran currículos de programas de humanidades, ciencias sociales y afines que atienden poco o nada su fomento y desarrollo, hecho que podría asociarse con el bajo desempeño académico en la competencia de razonamiento cuantitativo, tal como se encontró en esta investigación.

Analizando el modelo de la figura 38, en la competencia de inglés el buen desempeño se asocia al hecho de que la IES esté acreditada y que el estudiante no tenga personas a cargo. En el caso de que la IES no esté acreditada y la modalidad sea presencial, también constituye un patrón de buen desempeño el hecho de que los estudiantes pertenezcan a estratos altos (4, 5 y 6).

Por su parte, el bajo desempeño en inglés tiene relación con una IES no acreditada y con modalidad a distancia. En el caso de la modalidad presencial, el bajo desempeño se asocia con niveles de Sisben bajos (1 y 2) y con estratos socioeconómicos bajos (1 y 2). Estos hechos coinciden con Garbanzo (2007), Seibold (2000) y Montero y Villalobos (2004), en el sentido de que un resultado generalmente aceptable en el desempeño académico es la existencia de una asociación significativa entre el nivel socioeconómico del estudiante y su desempeño académico, hecho que coincide con lo encontrado en la presente investigación, en cuanto que forma parte de los patrones descubiertos, particularmente en la competencia genérica de inglés.

En resumen, entre los patrones de desempeño académico de las competencias genéricas, descubiertos y considerados en esta discusión, la acreditación institucional se encuentra asociada con otros atributos, tanto en el desempeño por encima de la media como por debajo de ella, excepto en lectura crítica, en la cual constituye en sí mismo un patrón de desempeño académico por encima de la media. Entre los atributos se destacan la modalidad del programa, el área del conocimiento, el número de instituciones en la zona, el número de estudiantes por zona, el estrato socioeconómico, el número de personas a cargo del estudiante y el género, entre otros; atributos que, en relación con el desempeño académico, de alguna manera, son considerados en los resultados de investigaciones de diversos autores.

Por otra parte, también se encontró que existe asociación para el buen desempeño académico en las competencias genéricas de lectura crítica y comunicación escrita, un alto número de estudiantes e instituciones por zona y otros atributos. En este contexto, conviene destacar que las variables denominadas número de IES y número de estudiantes por zona se crearon en el marco de esta investigación y se convierten en atributos importantes para el desarrollo de la educación superior en las zonas y regiones a las cuales pertenecen; por lo cual deberían ser consideradas en las políticas públicas de mejoramiento de la calidad en la formación de profesionales y su cobertura.

Así mismo, el área de los grupos de referencia de los programas académicos está asociada con el buen y mal desempeño en la competencia genérica de razonamiento cuantitativo. Otros atributos en los patrones descubiertos en esta investigación, asociados al buen y mal desempeño académico en las competencias de lectura crítica e inglés son el estrato socioeconómico y el nivel de Sisben al que pertenece el estudiante.

Finalmente, otro de los hallazgos de este estudio es la identificación del atributo género en los patrones de desempeño académico. En efecto, el género femenino se asocia, junto con otros atributos, al buen desempeño en la competencia de comunicación escrita, y el género masculino, con un bajo desempeño en dicha competencia; no obstante, se destaca que tales diferencias no son significativas.

Estos resultados no se pueden obtener a través de consultas simples a las bases de datos, ya que se desconocen las relaciones de los atributos con las clases. Únicamente, aplicando las técnicas y algoritmos de minería de datos, que en el caso de esta investigación es clasificación por árboles de decisión con el algoritmo J48, se logra obtener patrones asociados con el buen a mal rendimiento académico en las pruebas Saber Pro 2011-2.

Es necesario reconocer que la exactitud de los modelos de clasificación por árboles de decisión obtenidos para las cuatro competencias genéricas no es muy alta, pero sí satisfactoria para poder obtener en una primera aproximación, con esta técnica, los patrones que asocian los factores sociodemográficos, económicos, académicos e institucionales con el desempeño académico de los estudiantes que presentaron las pruebas Saber Pro 2011-2. Esto se debe a que se utilizaron como fuentes de datos únicamente las bases de datos del ICFES, en las cuales la llave principal para identificar cada estudiante es ficticia, lo que imposibilita el cruce de valores con otras fuentes de datos externas como las bases de datos del Sisben, de la Registraduría Nacional y del MEN, entre otras, que seguramente hubiesen mejorado la calidad de los datos y por ende la exactitud de los modelos.

De todas formas, se debe continuar investigando en este tema, seguramente identificando nuevos atributos o reevaluando los utilizados, mejorando las técnicas de limpieza y transformación utilizadas o usando otras técnicas de minería de datos que nos permitan mejores resultados en cuanto a precisión.