



Reglas de Asociación como Estrategia de Selección de Ítems en Exámenes Adaptativos Computarizados

Pacheco Ortiz, J., Rodríguez Mazahua, L. y Alor Hernández, G. (2022). Reglas de Asociación como Estrategia de Selección de Ítems en Exámenes Adaptativos Computarizados. *CIENCIA ergo-sum*, 30(2). *Postprint*.

Sección: Espacio del Divulgador

Universidad Autónoma del Estado de México, México

Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional.



Esta versión del artículo es una “versión final del autor” que fue aceptada por un proceso de **evaluación por pares ciegos**. Este documento diferirá en formato respecto a la “versión del editor”, la cual se someterá a un proceso de corrección de estilo y de diseño editorial. De ninguna forma se modificará el contenido. Todas las ideas que se presentan son responsabilidad del autor.

Reglas de Asociación como Estrategia de Selección de Ítems en Exámenes Adaptativos Computarizados

Association Rules as a Item Selection Strategy in Computerized Adaptive Testing

Josué Pacheco Ortiz, Tecnológico Nacional de México/I.T., México

Correo electrónico: iscjpo@gmail.com

ORCID: <https://orcid.org/0000-0002-2694-2827>

Lisbeth Rodríguez Mazahua, Tecnológico Nacional de México/I.T. Orizaba, México¹

Correo electrónico: lrodriguez@ito-depi.edu.mx

ORCID: <https://orcid.org/0000-0002-9861-3993>

Giner Alor Hernández, Tecnológico Nacional de México/I.T. Orizaba, México

Correo electrónico: giner.ah@orizaba.tecnm.mx

ORCID: <https://orcid.org/0000-0003-3296-0981>

Recepción: 16 de noviembre de 2021

Aprobación: 16 de marzo de 2022

RESUMEN

En el presente trabajo se propone un método de selección de ítems en los CAT (Computerized Adaptive Testing, Exámenes Adaptativos Computarizados) basado en reglas de asociación. De acuerdo con la respuesta correcta (1) o incorrecta (0) del evaluado a cada pregunta, se buscan reglas de asociación que tengan como antecedente la pregunta contestada y su resultado (1 o 0), y como consecuente, ítems con mayor o menor nivel de complejidad que el ítem respondido, respectivamente. La elección del ítem se realiza considerando las reglas con la mayor confianza y soporte. Finalmente, se incluye un ejemplo para demostrar la eficacia del método propuesto.

PALABRAS CLAVE: exámenes adaptativos computarizados, minería de datos, reglas de asociación.

ABSTRACT

This paper presents an item selection method in CAT (Computerized Adaptive Testing) based on association rules. According to the correct (1) or incorrect (0) answer of the examinee to each item, association rules are searched with the answered item and its result (1 or 0) as antecedent, and items with greater or lower complexity level than the responded item as a consequent, respectively. The election of the item is performed considering rules with the highest confidence and support. Finally, we include an example to demonstrate the efficacy of the proposed method.

KEYWORDS: association rules, computerized adaptive testing, data mining.

¹ Autor para correspondencia.

INTRODUCCIÓN

Los CAT son una forma de evaluación cuya característica principal es la de adaptar las preguntas siguientes de acuerdo con las respuestas otorgadas por el evaluado en las preguntas previas.

El algoritmo de selección de ítems dentro del proceso de los CAT inicia seleccionando la primera pregunta de dificultad media, y dependiendo de la respuesta que otorgue la persona evaluada, adaptará la dificultad de la siguiente pregunta en presentarse. Maximizar la precisión del examen es el objetivo principal de los CAT. Una de las ventajas principales de los CAT es que ocupan menos preguntas y toman menos tiempo que los exámenes tradicionales en lápiz y papel, conservando o mejorando el nivel de precisión (López-Cuadrado et al., 2010).

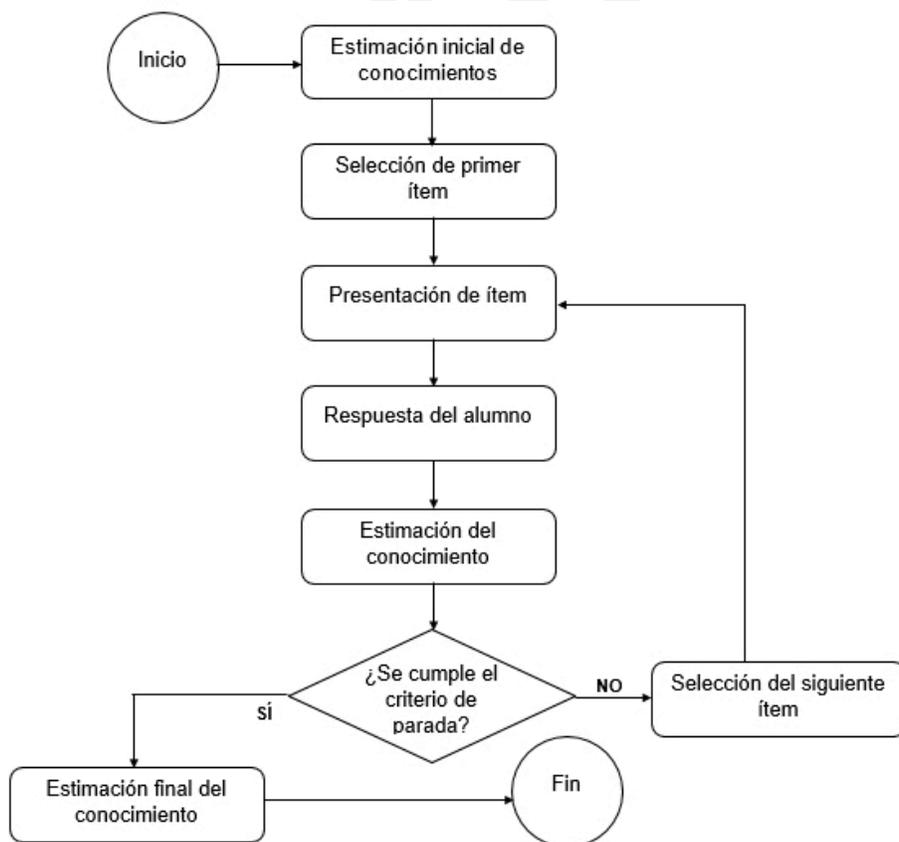
Uno de los componentes principales de los CAT es el criterio de selección de ítems. Actualmente la Información Máxima de Fisher (Albano et al., 2019) es uno de los métodos más empleados, sin embargo, presenta diversas desventajas, entre las cuales se encuentran: sesgo en la elección de ítems, error en la estimación de habilidad y duplicidad de ítems presentados (Sheng et al., 2018) (Lin & Chang, 2019) (Yigit et al., 2019), algunos trabajos se han enfocado en tratar de resolver los problemas presentados utilizando otras herramientas en la etapa de selección como Algoritmo Voraz (Bengs et al., 2018) y Kullback-Leibler Information (Tokusada & Hirose, 2016), sin embargo, a pesar de que han tenido resultados favorables, estos solo han sido en estudios de simulación y no en aplicación con grupos reales de evaluados.

En este artículo, se presenta un método de evaluación adaptativa computarizada que utiliza la minería de reglas de asociación como estrategia de selección de ítems, con lo cual se busca utilizar las ventajas potenciales de las reglas de asociación para encontrar relaciones entre las preguntas contestadas correcta o incorrectamente y las preguntas contestadas correctamente y presentar las preguntas más adecuadas (con mayor probabilidad de contestar correctamente) en los exámenes de acuerdo a las respuestas del evaluado, considerando las mejores reglas (guardadas en la base de datos de alumnos que contestaron el mismo examen previamente) con mayor soporte y confianza, es importante hacer mención que estas dos características son en las que se basa el algoritmo para hacer una correcta elección de la regla para la selección de la pregunta.

En la Figura 1 se muestra el proceso seguido por los CAT (Oppl et al., 2017), el cual comienza con la estimación inicial de conocimientos de la persona evaluada, posteriormente el primer ítem es seleccionado y presentado a la persona evaluada. Una vez que la respuesta es obtenida, una nueva estimación de conocimientos es llevada a cabo. La siguiente etapa consiste en verificar si el criterio de parada se cumple. Si esto no sucede, el siguiente ítem es seleccionado y el ciclo comienza de nuevo, el cual se repite hasta que se cumple el criterio de parada, cuando esto sucede, se lleva a cabo una estimación final de conocimiento y termina el proceso.

A lo largo del tiempo, diversos métodos, herramientas y estrategias han sido empleados en las distintas fases del CAT, por ejemplo: Modelo Logístico de Tres Parámetros (3PL) para la calibración de ítems; Estimación de Máxima Verosimilitud (MLE) para la estimación de habilidad del evaluado; y Diferencias Cuadráticas Medias como criterio de evaluación.

Figura 1
Proceso CAT



Fuente: Elaboración propia.

1. REGLAS DE ASOCIACIÓN

La minería de reglas de asociación tiene como finalidad encontrar relaciones entre los elementos u objetos de una base de datos, es decir, permite descubrir patrones comunes en los elementos; se ha utilizado en diversas áreas, entre las cuales se encuentran: salud (Zhang et al., 2020) y aprendizaje en línea (Paladines Rodríguez et al., 2021), ofreciendo resultados positivos en cada uno de los casos.

Existen diversos algoritmos para la minería de reglas de asociación, entre los cuales destacan los siguientes: Apriori (Agrawal et al., 1993), FPGrowth (Han et al., 2016), PredictiveApriori (Scheffer, 2001) y Tertius (Flach & Lachiche, 2001). Además, para el análisis comparativo de los algoritmos de reglas de asociación, se utilizan diversos criterios, entre los cuales destacan los siguientes:

- a) Confianza: evalúa el grado de certeza de la asociación detectada.
- b) Soporte: representa el porcentaje de transacciones de la base de datos que satisface la regla dada.
- c) Tiempo: número de milisegundos que toma la construcción de un modelo.
- d) Reglas: representa el número de reglas interesantes obtenidas.
- e) Sustentación: indica la relación entre el soporte observado de un conjunto de elementos y el soporte teórico de ese conjunto dada la suposición de independencia.

2. PROCESO CAT CON REGLAS DE ASOCIACIÓN

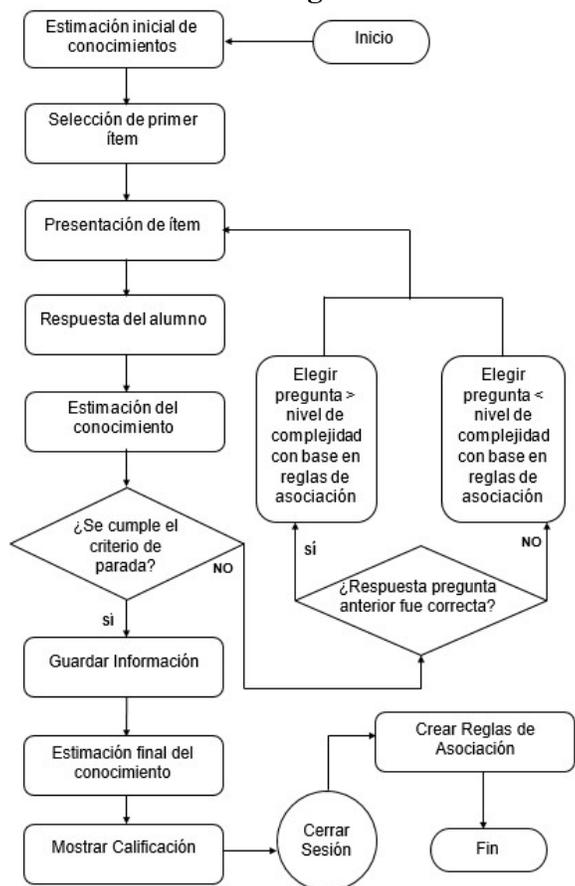
La Figura 2 muestra el método de evaluación adaptativa computarizada propuesto que integra a la minería de reglas de asociación como estrategia de selección de ítems en el proceso de los CAT, las cuales se emplean de la siguiente forma:

- a) Basándose en que cada pregunta o ítem tiene un nivel de complejidad (1 a 100), el sistema elige la pregunta inicial de acuerdo con el criterio de arranque seleccionado, el cual es la elección de una pregunta de nivel medio.
- b) Para la elección de la siguiente pregunta:

- Si la respuesta de la pregunta inmediata anterior fue correcta ($pX=1$), el sistema elige una de mayor nivel de complejidad (pY) tomando en cuenta las reglas de asociación almacenadas en la base de datos (de alumnos que presentaron la misma prueba previamente), con mayor soporte y confianza generadas de la siguiente forma:

$$pX=1 \Rightarrow pY=1.$$

Figura 2
Proceso CAT con Reglas de Asociación



Fuente: Elaboración propia.

- Si la respuesta de la pregunta inmediata anterior fue errónea ($pX=0$), el sistema elige una de menor nivel de complejidad (pZ) tomando en cuenta las reglas de asociación almacenadas en la base de datos (de alumnos que presentaron la misma prueba previamente), con mayor soporte y confianza generadas de la siguiente forma:

$$pX=0 \Rightarrow pZ=1.$$

c) El proceso del paso 2 se repite con cada una de las preguntas siguientes.

4 Después de finalizar el número de preguntas establecidas por el profesor, el sistema termina el examen, guarda toda la información generada en su base de datos y proporciona la calificación al alumno.

d) El alumno cierra sesión o regresa a la página inicial para presentar otro examen.

Finalmente, el sistema genera reglas de asociación a través de la aplicación inmediata del algoritmo Apriori (Agrawal et al., 1993) a la información guardada, con la finalidad de encontrar relaciones entre las preguntas contestadas y que sirvan como base para la próxima vez que un alumno presente el mismo examen.

3. SELECCIÓN DE ÍTEMS CON REGLAS DE ASOCIACIÓN

La Figura 3 muestra el proceso completo de la fase de selección de ítems aplicando reglas de asociación, misma que se describe a continuación:

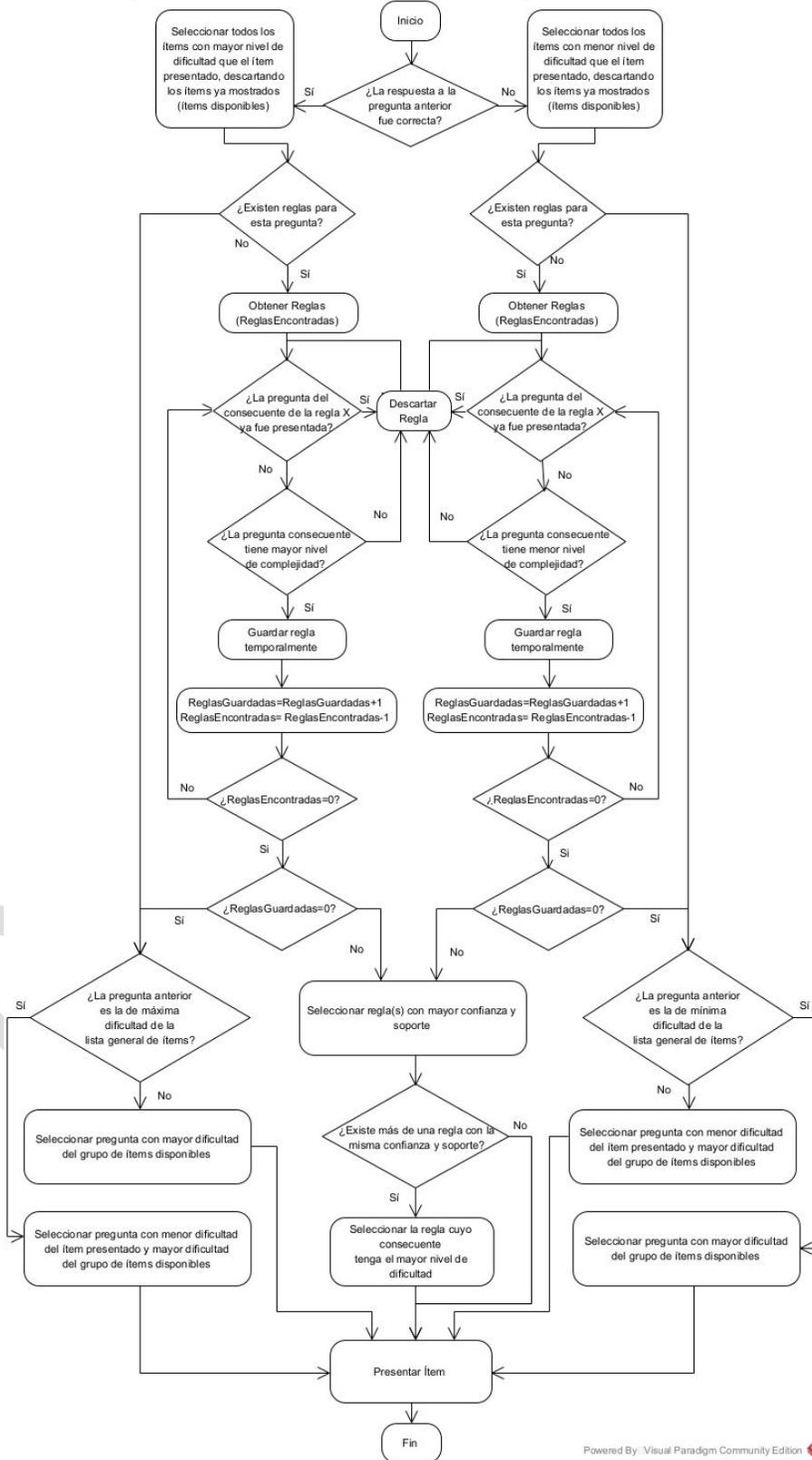
El proceso de la fase de selección de ítems inicia cuando el algoritmo verifica si la respuesta a la pregunta anterior fue correcta o incorrecta, en caso de ser correcta:

a) Se seleccionan todos los ítems con mayor nivel de dificultad que el ítem presentado, descartando los ítems ya mostrados anteriormente. A este conjunto de ítems se le denomina “ítems disponibles”.

b) Se verifica si existen reglas de asociación almacenadas en la base de datos para el ítem presentado. En caso de existir, se obtienen y a ese conjunto se le denomina “reglas encontradas”, en caso de no existir reglas, se procede al Paso 5.

c) El siguiente paso consiste en revisar cada una de las reglas encontradas: primero se revisa si la pregunta consecuente de la regla ya fue presentada previamente al alumno, en caso de ser afirmativo se descarta la regla. Si no ha sido presentada, entonces ahora se verifica si la pregunta consecuente tiene mayor nivel de complejidad que la pregunta presentada previamente, en caso de ser negativo se descarta la regla. Si la regla ha pasado los dos filtros previos, entonces se guarda de manera temporal en un grupo llamado “reglas guardadas”.

Figura 3
Etapa de Selección de ítems con Reglas de Asociación



Fuente: Elaboración propia.

d) El paso 3 se repite para cada una de las “reglas encontradas”, cuando ya no hay más, entonces se comprueba si el conjunto de “reglas guardadas” es igual a cero, en caso de ser afirmativo se procede al paso 5. En caso de ser negativo, se seleccionan reglas con mayor confianza y soporte, se procede a verificar si existe más de una regla con la misma confianza y soporte, en caso de ser negativo se procede al paso 6, en caso de ser positivo, se selecciona la regla cuyo consecuente tenga el mayor nivel de dificultad y se procede al paso 6.

e) Se verifica si la pregunta recién presentada es la de máxima dificultad en la lista general de ítems, en caso de ser afirmativo, se selecciona una pregunta con menor dificultad del ítem presentado y mayor dificultad del grupo de ítems disponibles, se procede al paso 6. En caso de ser negativo, se selecciona la pregunta con mayor dificultad del grupo de ítems disponibles, se procede al paso 6.

f) Se presenta la pregunta y termina la etapa de selección de ítems.

Para el caso de que la respuesta a la pregunta anterior sea incorrecta, prácticamente se sigue el mismo proceso, con las siguientes diferencias:

- En el Paso 1, se seleccionan ítems con menor nivel de dificultad que el ítem presentado.
- En el Paso 3, se verifica si la pregunta consecuente de la regla tiene menor nivel de complejidad.
- En el Paso 5, se verifica si la pregunta reciente presentada es la de mínima dificultad en la lista general de ítems, en caso de ser afirmativo, se selecciona la pregunta con mayor nivel de dificultad del grupo de ítems disponibles. En caso de ser negativo, se selecciona una pregunta con menor dificultad del ítem presentado y mayor dificultad del grupo de ítems disponibles, en ambos casos igual se procede al Paso 6.

3. 1. Ejemplo

Suponiendo que se toma el mismo nivel de estimación inicial de habilidad para todos los alumnos de un grupo (pregunta con dificultad media) y se tienen 10 preguntas identificadas

como Pregunta1 a Pregunta10, ordenadas de acuerdo con su nivel de dificultad de menor a mayor respectivamente; la última pregunta presentada fue la Pregunta7; y la respuesta del alumno fue correcta, entonces:

a) El algoritmo selecciona todas las preguntas con mayor nivel de dificultad que la pregunta presentada, descartando las preguntas ya presentadas: Pregunta8, Pregunta9 y Pregunta10 (Paso 1).

b) Se verifica si existen reglas para la Pregunta7, teniendo dos:

- Pregunta7(1) => Pregunta8(1) Confianza=1, Soporte 0.90
- Pregunta7(1) => Pregunta10(1) Confianza=1, Soporte 0.90

Lo cual significa que si se contesta bien la Pregunta7, entonces implica mayor probabilidad de responder correctamente la Pregunta8 y la Pregunta10. Este conjunto de reglas reciben el nombre de “reglas encontradas” (Paso 2).

c) El algoritmo comprueba si la pregunta consecuente (Pregunta8) de la primera regla encontrada ya fue presentada, en este caso no ha sido presentada, entonces procede a verificar si la pregunta tiene mayor nivel de dificultad, en este caso sí presenta mayor nivel de dificultad, por lo tanto, procede a almacenarla de manera temporal en “reglas guardadas”. Se repite el proceso con la pregunta consecuente (Pregunta10) de la segunda regla, siendo el mismo resultado y también la almacena (Paso 3).

d) El siguiente paso en el algoritmo consiste en verificar si existe más de una regla en “reglas guardadas” que tenga el mismo nivel de confianza y soporte, para este ejemplo, ambas reglas tienen los mismos niveles, por lo tanto, el algoritmo selecciona la regla cuya pregunta consecuente tenga el mayor nivel de dificultad, para este caso, la segunda regla tiene a la Pregunta10 que, de acuerdo con las bases del ejemplo, es la que tiene mayor nivel de dificultad (Paso 4).

e) Por último, la Pregunta10 es presentada al alumno (Paso 6).

PROSPECTIVA

La contribución principal de este trabajo es la propuesta de un método de evaluación adaptativa computarizada que integra a la minería de reglas de asociación como estrategia de selección de ítems. La principal ventaja de este método es que la selección se basará en los datos recolectados de otros estudiantes que han realizado el mismo examen con el objetivo de presentar al evaluado las preguntas con mayor probabilidad de responder correctamente. En el futuro, se desarrollará una aplicación web que implemente el método propuesto y se aplicará en un curso de Bases de Datos a nivel maestría para evaluar su desempeño.

CONCLUSIONES

Evaluar es una parte fundamental dentro del proceso enseñanza-aprendizaje, ya que no solamente le da una nota o calificación al alumno, sino que, además, le brinda información al profesor acerca de las ventajas y desventajas que tiene cada evaluado, lo cual puede ser utilizado para mejorar la calidad del proceso de enseñanza. A lo largo del tiempo, las pruebas o exámenes han sido la forma más de común de medir el nivel de conocimientos de las personas.

Mientras que en el sistema tradicional suele aplicarse el mismo examen para todas las personas que forman parte de un grupo físico, en el sistema e-Learning las pruebas son de manera digital y de la misma forma que en el sistema tradicional de enseñanza igual puede aplicarse el mismo examen para todos los alumnos que forman parte de un grupo virtual. Sin embargo, con el surgimiento de los CAT, las pruebas pueden adaptarse al nivel de conocimiento de los evaluados.

El proceso de los CAT contempla varias etapas, una de las principales es la selección de ítems, donde la Información Máxima de Fisher es uno de los criterios más empleados, aunque presenta ciertas desventajas tales como repetición de ítems presentados, sesgo en la elección, entre otras. Para tratar de resolver estos errores, el presente trabajo muestra la aplicación de la minería de reglas de asociación como una estrategia de selección de ítems, donde, basándose en el conocimiento generado en información de exámenes previamente contestados por alumnos, el algoritmo de selección puede elegir los ítems que mejor se adapten al alumno de acuerdo con su nivel de conocimiento.

El proceso de selección de ítems presentado se empleará en la implementación de un sistema de evaluación CAT, cuya característica principal es la de utilizar reglas de asociación en la etapa de selección de ítems, por el momento, el sistema de evaluación se encuentra en la etapa de desarrollo, al finalizar, se llevarán a cabo pruebas de simulación que permitirán verificar la eficacia del mismo.

AGRADECIMIENTOS

Los autores agradecen al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca de doctorado otorgada al primer autor, al Tecnológico Nacional de México (TecNM) por apoyar el desarrollo de este trabajo y a los árbitros de la revista por sus aportes que permitieron mejorar el contenido de este artículo.

REFERENCIAS

Agrawal, R., Imielinski, T., & Swam, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 22(2), 207-216.

Albano, A., Cai, L., Lease, E., & McConnell, S. (2019). Computerized Adaptive Testing in Early Education: Exploring the Impact of Item Position Effects on Ability Estimation. *Journal of Educational Measurement*, 56(2), 437-451.
<https://doi.org/https://doi.org/10.1111/jedm.12215>

Bengs, D., Brefeld, U., & Kröhne, U. (2018). Adaptive Item Selection Under Matroid Constraints. *Journal of Computerized Adaptive Testing*, 6(2), 15-36.
<https://doi.org/http://dx.doi.org/10.7333/1808-0602015>

Flach, P., & Lachiche, N. (2001). Confirmation-Guided Discovery of First-Order Rules with Tertius. *Machine Learning*, 42(1-2), 61-95.

Han, J., Kamber, M., & Pei, J. (2016). *Data Mining: Concepts and Techniques* (3° ed.). Morgan Kaufmann.

Lin, C.-J., & Chang, H.-H. (2019). Item Selection Criteria With Practical Constraints in Cognitive Diagnostic Computerized Adaptive Testing. *Educational and Psychological*, 79(2), 1-23. <https://doi.org/https://doi.org/10.1177/0013164418790634>

López-Cuadrado, J., Pérez, T., Vadillo, J., & Gutiérrez, J. (2010). Calibration of an item bank for the assessment of Basque language knowledge. *Computers & Education*, 55, 1044-1055.

Oppl, S., Reisinger, F., Eckmaier, A., & Helm, C. (2017). A flexible online platform for computerized adaptive testing. *International Journal of Educational Technology in Higher Education*, 14(2), 2-21. <https://doi.org/https://doi.org/10.1186/s41239-017-0039-0>

Paladines Rodríguez, J., Carreño Arce, C., & Parrales Loor, A. (2021). Incidencia en la deserción de los cursos de primer nivel de la carrera de desarrollo de software del instituto superior tecnológico Guayaquil. *Serie Científica de la Universidad de las Ciencias Informáticas*, 14(6), 43-58.

Scheffer, T. (2001). Finding association rules that trade support optimally against confidence. *Principles of Data Mining and Knowledge Discovery*, 9(4), 224-235.

Sheng, C., Bingwei, B., & Jiecheng, Z. (2018). An Adaptive Online Learning Testing System. *ICIET '18 Proceedings of the 6th International Conference on Information and Education Technology*, 55(1), 18-24.

<https://doi.org/https://doi.org/10.1145/3178158.3178187>

Tokusada, Y., & Hirose, H. (2016). Evaluation of Abilities by Grouping for Small IRT Testing Systems. *5th IIAI International Congress on Advanced Applied Informatics*, 445-449. <https://doi.org/https://doi.org/10.1109/IIAI-AAI.2016.50>

Yigit, H., Sorrel, M., & de la Torre, J. (2019). Computerized Adaptive Testing for Cognitively Based Multiple-Choice Data. *Applied Psychological Measurement*, 43(5), 1-14.

Zhang, D., Jintao, L., Zhang, B., Zhang, X., Jiang, H., & Lin, Z. (2020). The characteristics and regularities of cardiac adverse drug reactions induced by Chinese materia medica: A bibliometric research and association rules analysis. *Journal of Ethnopharmacology*, 252, 1-12.

CC BY-NC-ND