

Hacia el desarrollo de un sistema de evaluación adaptativa computarizada basada en reglas de asociación secuenciales

Towards the development of a computerized adaptive testing system based on sequential association rules

A. Reyes-García ^{a, *}, L. Rodríguez-Mazahua ^a, J. Pacheco-Ortiz ^a, M. Abud-Figueroa ^a, U. Juárez-Martínez ^a

^a Tecnológico Nacional de México, Instituto Tecnológico de Orizaba, Orizaba, Ver., México.

Resumen

En este trabajo se muestran los resultados de un análisis de 38 artículos de investigación de las principales editoriales informáticas, para comparar los métodos de minería de patrones secuenciales que utilizaron, además se propone un sistema de evaluación adaptativa computarizada (CAT) basada en reglas de asociación secuenciales. Anteriormente, se presentó un sistema CAT que integra reglas de asociación como un método de selección de ítems. En este caso, el sistema suministrará a los estudiantes ítems que correspondan a su nivel de conocimiento a través de las repuestas que ellos ingresen en la prueba, además de las respuestas correctas de otros estudiantes en la misma prueba basadas en reglas de asociación secuenciales con mayor soporte. Las tecnologías propuestas para el desarrollo del sistema son Java, y el entorno de desarrollo NetBeans, además del marco de trabajo JavaServer Faces, el sistema gestor de bases de datos MySQL, apegado a la metodología de ingeniería web basada en UML (UWE por sus siglas en inglés).

Palabras Clave: Exámenes Adaptativos Computarizados, Minería de datos, Reglas de Asociación Secuenciales, Educación.

Abstract

This paper shows the results of an analysis of 38 research articles from the main Computer Science publishers, to compare the sequential pattern mining methods they used, additionally, a computerized adaptive testing system (CAT) based on sequential association rules is proposed. Previously, a CAT system integrating association rules as an item selection method was proposed. In this case, the system will provide the students with items that correspond to their level of knowledge through the answers that they enter the test, in addition to the correct answers of other students in the same test based on sequential association rules with greater support. The technologies proposed for the development of the system are Java, and the integrated development environment NetBeans, along with the JavaServer Faces framework, the MySQL database management system, following the UML-based web engineering methodology (UWE for its acronym in English).

Keywords: Computerized Adaptive Testing, Data Mining, Sequential Association Rules, Education.

1. Introducción

Un examen o una evaluación es un instrumento por el cual netamente se evalúa el nivel de aprendizaje y conocimiento que tiene un sujeto o estudiante que lo presenta, estas evaluaciones son consideradas como herramientas psicopedagógicas universales en el contexto escolar (Salmerón, 2011). Los exámenes tradicionales realizados a papel cuentan con una serie de preguntas que son de tipo

abierto o cerrado. Sin embargo, estas evaluaciones suponen que todos los estudiantes cuentan con el mismo nivel de habilidad y conocimiento, por eso los incisos se plantean con el mismo nivel de dificultad para todos (Istiyono et al., 2020). Por lo consiguiente con el surgimiento de la tecnología, la forma de evaluar a los alumnos cambió de tal forma que se realizaron diferentes maneras de aplicar exámenes como a través de plataformas, es ahí donde surgen los CAT (exámenes adaptativos computarizados) que consisten en nivelar las

*Autor para la correspondencia: msc.areyes@ito-depi.edu.mx

Correo electrónico: msc.areyes@ito-depi.edu.mx (Arturo Uriel Reyes-García), Irodruiguez@ito-depi.edu.mx (Lisbeth Rodríguez-Mazahua), iscjpo@gmail.com (Josué Pacheco-Ortiz), maria.af@orizaba.tecnm.mx (María Antonieta Abud-Figueroa), ulises.jm@orizaba.tecnm.mx (Ulises Juárez-Martínez).

preguntas de la prueba con el nivel de conocimiento del estudiante, donde el componente más importante es el método de selección de ítems (Weiss y Kingsbury, 2016). De manera que en los últimos años se realizaron diversas propuestas para un mejor método de selección de ítems donde destaca una estrategia utilizando reglas de asociación (Pacheco-Ortiz et al., 2021), esta estrategia selecciona los ítems de acuerdo con la respuesta correcta o incorrecta del evaluado y una regla de asociación con mayor confianza y soporte. Sin embargo, actualmente no se tiene registro de un CAT que utilice reglas de asociación secuenciales para este propósito. Debido a esta problemática, en este artículo se analizaron diversos trabajos de las principales editoriales informáticas para responder la pregunta ¿cuáles son las técnicas de minería de patrones secuenciales más utilizadas en aplicaciones educativas? También se propone el diseño de un sistema de evaluación adaptativa que integre reglas de asociación secuenciales como método de selección de ítems. El presente documento está organizado de la siguiente manera: la Sección 2 expone la metodología de análisis de los 38 trabajos que se encontraron y la tabla comparativa de los mismos. En la Sección 3 se presenta una propuesta de flujo de trabajo y la arquitectura de un sistema de evaluación adaptativa computarizada y en la Sección 4 se presentan la conclusión y trabajos a futuro.

2. Metodología de análisis

En la Figura 1 se muestra la metodología que se empleó para la búsqueda, análisis y revisión sistemática de un total de 38 artículos científicos en las bases de datos en línea de las principales bibliotecas digitales de Ciencias de la computación, por ejemplo: *IEEE Xplore*, *ACM digital library*, *SpringerLink*, *ScienceDirect*, y otras. Las palabras clave que se utilizaron son: *Sequential Pattern Mining and Education*, durante la selección se descartaron aquellos artículos escritos en español o que formaban parte de trabajos de tesis, además

solo se consideraron los que fueron publicados en el rango de los años 2017 al 2022. Posteriormente se realizó un análisis comparativo usando ocho criterios: 1) los algoritmos utilizados, 2) las tecnologías empleadas, 3) el área de estudio en la que se centró cada trabajo (Informática, Idiomas, entre otras), 4) el nivel de estudio considerado (primaria, secundaria y superior), 5) los conjuntos de datos usados, 6) el problema abordado, 7) el año de publicación, y 8) la editorial.

En la Tabla 1 se muestra un resumen del análisis comparativo de los trabajos relacionados considerando los cinco que se consideran más relevantes para este trabajo: 1) Evaluación: Se observa que solo en dos trabajos se han aplicado métodos de minería de reglas de asociación secuenciales para evaluar a los estudiantes: Mudrick et al. (2019) y Cheng et al. (2021); 2) Conjunto de datos real: La mayoría (74%) de los artículos revisados emplearon conjuntos de datos reales; 3) Datos universitarios: El 66% de los métodos revisados utilizaron datos de universidades para validar su estudio; 4) Herramienta de minería de datos: Solo dos trabajos mencionan que usaron WEKA (*Waikato Environment for Knowledge Analysis*, Entorno de Waikato para el Análisis de Conocimiento) para implementar sus métodos: Lien et al. (2020) y Yildirim y Usluel (2022); 5) Informática: El 13% de los artículos revisados se enfocaron en un área de Ciencias de la Computación diferente a Programación.

En este artículo, se presenta el flujo de trabajo y la arquitectura de un sistema CAT basado en reglas de asociación secuenciales que será validado utilizando datos de exámenes de la materia Bases de Datos de una universidad mexicana. Se propone utilizar WEKA para obtener las reglas de asociación secuenciales. Por lo tanto, el sistema presentado cumple con todas las características de este análisis.

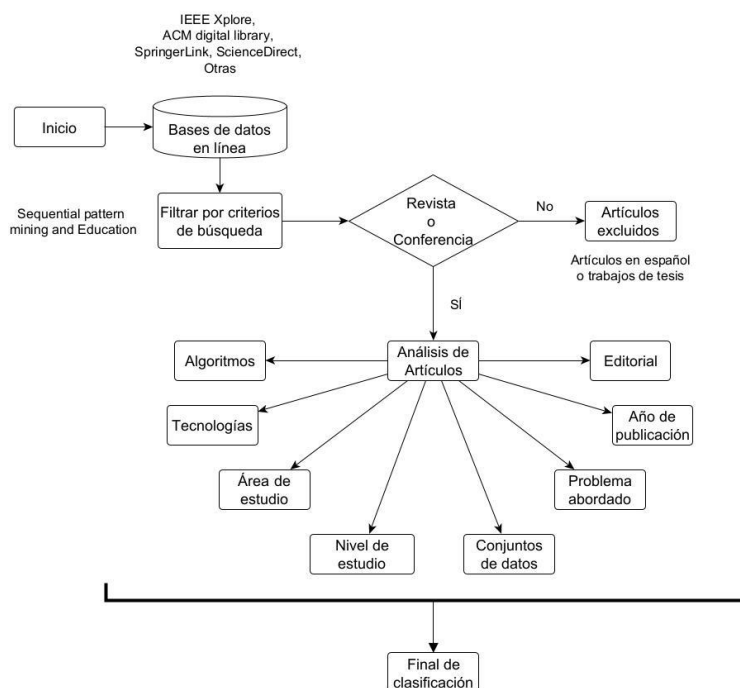


Figura 1: Criterios de selección.

Tabla 1: Trabajos relacionados.

Artículos	1	2	3	4	5
(Tarus et al., 2018)		X	X		
(Anwar y Uma, 2022)		X			
(Zheng et al., 2019)		X	X		
(Wong et al., 2019)		X	X		
(Liu et al., 2018)	X	X	X		
(Qu et al., 2019)		X	X		
(Klašnja-Milićević et al., 2018)		X	X		
(Zhu et al., 2019)					
(Deeva y De Weerd, 2019)					
(Taub et al., 2018)		X	X		
(Wan y Niu, 2020)		X	X		
(Al-Twijri et al., 2022)					
(Taub y Azevedo, 2018b)		X	X		
(Norm Lien et al., 2020)		X		X	
(Yildirim y Usluel, 2022)		X	X	X	
(Zhang et al., 2022)		X	X		X
(Shih, 2018)		X	X		
(Song et al., 2022)		X	X		
(Wang y Zaiane, 2018)					
(Mudrick et al., 2018)		X	X		
(Taub y Azevedo, 2018a)		X	X		
(Yang, 2021)					
(Bermudez et al., 2020)					
(Cheng Tan et al., 2020)		X	X		
(Malekian et al., 2020)		X	X		
(Kong y Pollock, 2020)		X	X		
(Latypova, 2022)		X	X		X
(Fatahi et al., 2018)		X	X		X
(Pogorskiy y Beckmann, 2022)		X			
(He et al., 2021)		X			
(Chen y Wang, 2020)		X			
(Czibula et al., 2019)		X	X		
(Doko et al., 2018)		X			X
(Real et al., 2021)		X			
(Song y Ye, 2021)					
(Cheng et al., 2021)	X				
(Niemeijer et al., 2020)					X
(Aktas y Aktas, 2021)					

2.1. Clasificación de los trabajos por año de publicación y editorial

En la Figura 2 se muestra la clasificación de los artículos por el año en que se publicaron, estos se encuentran entre el año 2017 y el año 2022. Además, se expone que el mayor número de estos se publicó en 2018, los trabajos publicados en este año abordaron diferentes problemáticas como, por ejemplo, en Liu et al. (2018) se enfrentó el problema de la extracción de características semánticas en preguntas en enseñanza del idioma inglés, por su parte, Klašnja-Milićević et al. (2018) aportó un enfoque híbrido para la recomendación de recursos educativos personalizados, también en Taub et al. (2018) se estudiaron las causas de un mal aprendizaje autorregulado, mientras que en Taub y Azevedo. (2018b) se estudió la falta de estrategias para la implementación de este. Además, en Shih. (2018) también se abordó el tema de falta de

métodos de aprendizaje efectivos en estudiantes de programación. De la misma forma, Wang y Zaiane. (2018) trató el tema de una correcta recomendación de recursos educativos, al igual que Mudrick et al. (2018) quien estudio el monitoreo metacognitivo para el desarrollo de un sistema de recomendación de cursos, por último, en Doko et al. (2018) se recomendó un modelo y la arquitectura de un sistema para la visualización de videos educativos.

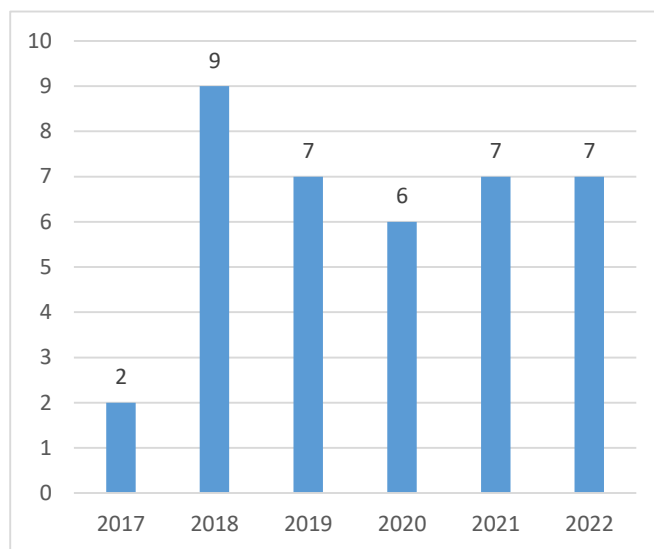


Figura 2: Artículos por año de publicación.

En la Figura 3 se observa el número de artículos que se obtuvieron de cada editorial, resaltan las editoriales Springer y Elsevier de donde se obtuvieron el mayor número de trabajos. Dentro de los trabajos que destacan están los de Zheng et al. (2019), Taub et al. (2018), He et al. (2021) y Zhu et al. (2019), ya que su área de estudio fue STEM (*Science, Technology, Engineering, and Mathematics*, Ciencia, Tecnología, Ingeniería y Matemáticas), el resto de los trabajos se centraron en áreas como sociales (Wong et al., 2019) o biología (Mudrick et al., 2018), sin embargo, hubo algunos trabajos que no mencionaron su área de estudio, este es el caso de Anwar y Uma. (2022), Pogorskiy y Beckmann. (2022) y Czibula et al. (2019) pertenecientes a Elsevier. En la clasificación de los trabajos de Springer, se encuentra el trabajo de Song et al.(2022) en el que se realiza la propuesta de dos nuevos algoritmos para la extracción de patrones secuenciales, al igual que en Al-Twijri et al. (2022) donde se propone otro nuevo algoritmo, además de esto en el trabajo antes mencionado y Chen y Wang. (2020) se hace uso del SPMF (*sequential pattern mining framework*, marco de minera de patrones secuenciales) como herramienta de minería de datos, por otro lado, Deeva y De Weerd, (2019) y Zhang et al.(2022) se enfocaron en entornos de aprendizaje, mientras que el resto como Taub y Azevedo. (2018b) se centró en áreas de interés para el aprendizaje autorregulado y Wang y Zaiane. (2018) en la recomendación de recursos.

2.2. Problema abordado

Para la siguiente clasificación se identificaron siete temáticas de problemas abordados en los trabajos relacionados donde destaca el tema de predicción de comportamiento al tener más artículos que abordan esta problemática, los resultados de los demás temas se muestran en la Figura 4.

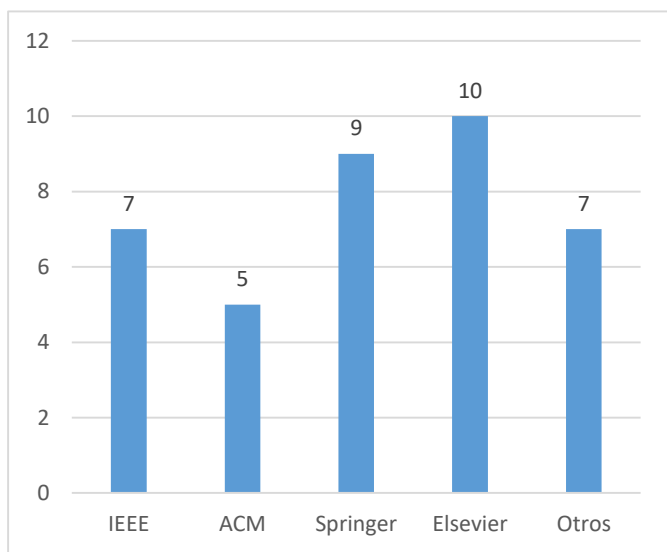


Figura 3: Artículos por editorial.

2.3. Conjuntos de datos

Se realizó un análisis acerca de si los conjuntos de datos usados en los trabajos relacionados eran sintéticos o reales, sin embargo, el 74% describe cuál es la procedencia de su conjunto siendo catalogados como de tipo real y el 27% no menciona ninguna descripción de sus conjuntos, además de que también se analizó si los conjuntos reales se encontraban disponibles para su descarga y se encontró que solo los trabajos de Anwar y Uma. (2022), Song et al.(2022) y He et al. (2021) están públicamente disponibles, dichos resultados se muestran en la Figura 5.

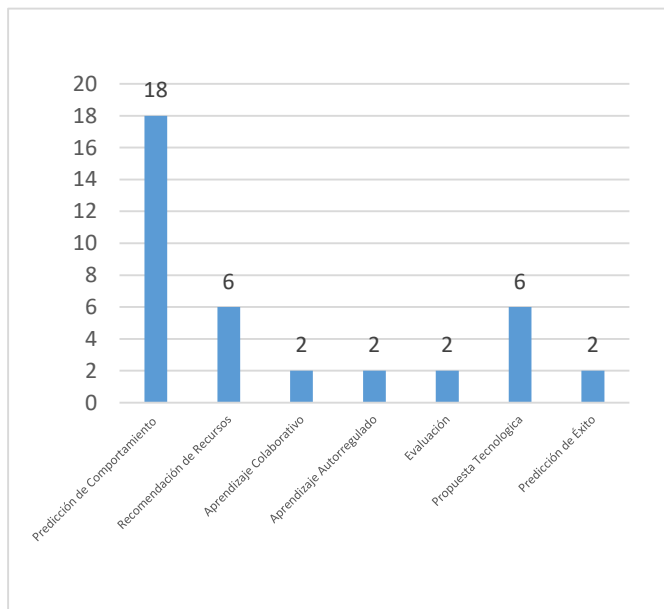


Figura 4: Grafica de problemas abordados.

2.4. Nivel de estudios

En este criterio se observó el nivel de estudio al cual pertenecen los participantes de cada conjunto de datos

identificado en los trabajos relacionados, cabe resaltar que en este rubro hubo conjuntos con datos conformados por dos tipos de nivel de estudios como en el caso de Zheng et al. (2019) que analizó datos de participantes con nivel de estudios de secundaria y superior, también resalta el trabajo de Norm Lien et al. (2020) que fue el único que usó datos provenientes de participantes de nivel primaria.

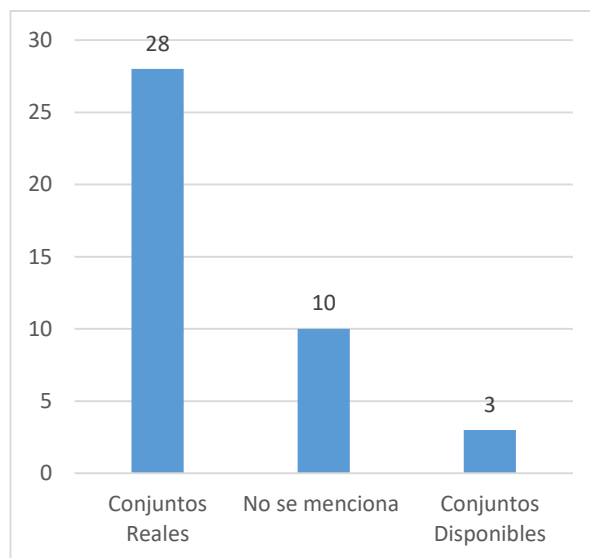


Figura 5: Tipo de conjunto de datos.

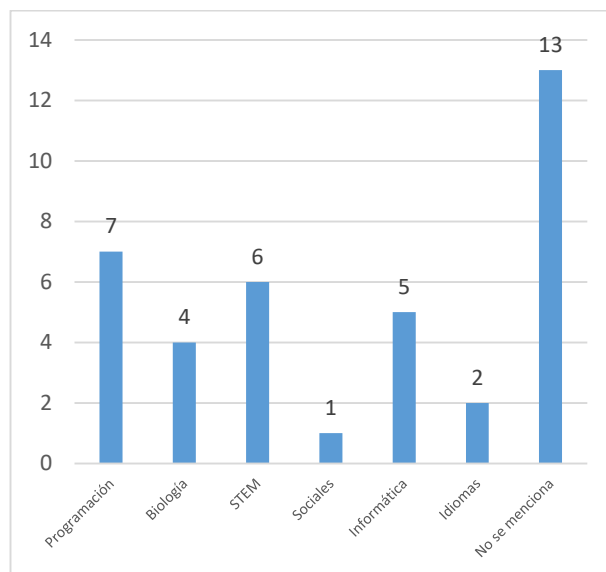


Figura 6: Área de estudio.

2.5. Algoritmos

En la Figura 8 se observa el número de algoritmos que se utilizaron para minería de patrones secuenciales en los trabajos relacionados, también es importante mencionar, que algunos de los trabajos usan más de un solo algoritmo como en el caso de Anwar y Uma (2022) que utilizó los algoritmos PrefixSpan y Topseq, además Aktas y Aktas (2021) utilizaron los algoritmos GSP, PrefixSpan y CMRules.

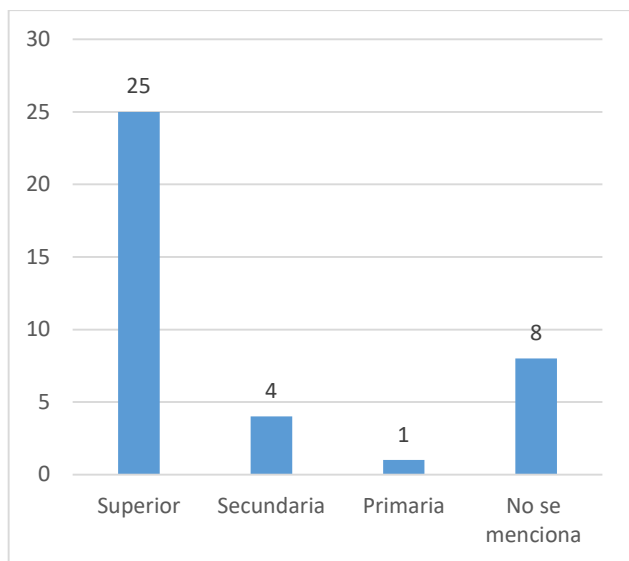


Figura 7: Nivel de estudio de los participantes.

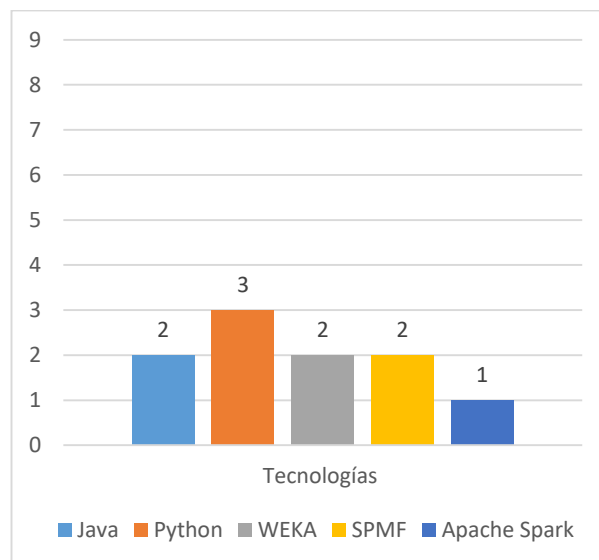


Figura 9: artículos por Tecnologías.

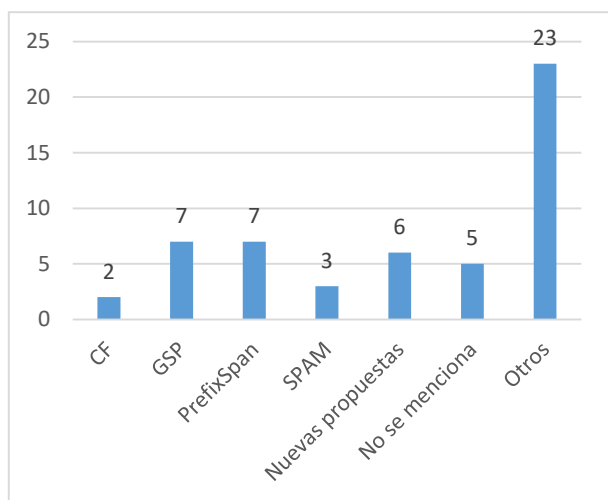


Figura 8: Artículos por algoritmos utilizados.

Después del análisis realizado se encontró que el mayor número de trabajos se publicó en el año 2018, además de esto la mayoría de estos fueron publicados principalmente en las editoriales Springer y Elsevier, otro de los resultados obtenidos fue que el problema principal abordado fue la predicción de comportamiento con un 47% de los trabajos usando técnicas de minería de patrones secuenciales para su resolución donde destacan los métodos de GPS (*Generalized Sequential Pattern*, Patrón Secuencial Generalizado) y PrefixSpan, sin embargo, muy pocos abordan el problema de evaluación y la mayoría usó como área de estudio la programación, también se usaron datos de estudiantes de nivel superior donde se comprobó que pocos estaban disponibles al público, por último, las tecnologías más usadas fueron las herramientas de minería de datos SPMF y WEKA, además de los lenguajes de programación Java y Python.

3. Sistema CAT propuesto basado en reglas de asociación secuenciales

En esta sección, se presenta la Figura 10 como propuesta de flujo de trabajo para el desarrollo del sistema de evaluación adaptativa computarizada basada en reglas de asociación secuenciales. El sistema propuesto logrará que cada elemento de la evaluación sea apropiado para el nivel de capacidad del sujeto examinado. El sistema CAT contará con un grupo de elementos calibrados con anterioridad y un punto de partida, para que el algoritmo de selección de ítems basado en reglas de asociación secuenciales determine cuál es el ítem que se administrará, donde la selección de la pregunta dependerá de las respuestas de otros estudiantes y de las respuestas anteriores del mismo sujeto evaluado. A continuación, en la Subsección 3.1 se explican los componentes más importantes del CAT. Mientras que en las Subsección 3.2 se expone la arquitectura propuesta para el sistema y se explican brevemente las tecnologías con las que se pretende desarrollar el mismo.

3.1. Componentes del CAT

2.6. Tecnologías

En la Figura 9 se muestran las tecnologías que se utilizaron en los trabajos relacionados, sin embargo, la mayoría de los trabajos no mencionan esta información, por eso en los resultados se observan pocas tecnologías, donde resalta el uso de los lenguajes de programación Java en Yildirim y Usluel (2022) y Yang (2021), Python en Shih (2018), Real et al. (2021) y Aktas y Aktas (2021), además de que se mencionan las herramientas de minería de datos WEKA en los trabajos de Norm Lien et al. (2020) y Yildirim y Usluel (2022), *Sequential Patter Mining Framework* (SPMF) en Shih (2018) y Chen y Wang (2020), por último, la herramienta *Apache Spark MLLib library*.

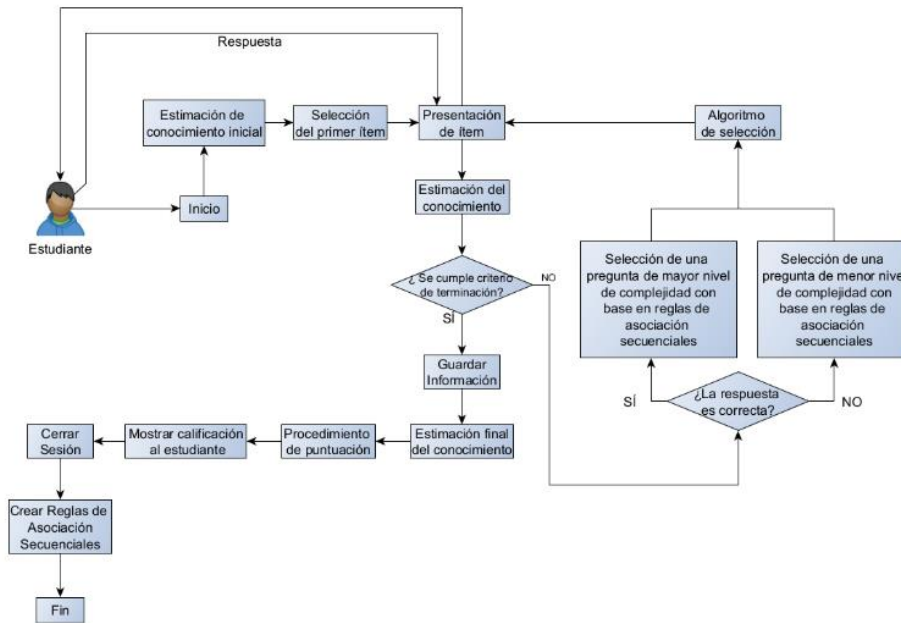


Figura 10: Flujo de trabajo de CAT.

Grupo de elementos calibrados: Este componente consiste en un conjunto de ítems disponibles para que el CAT elija. Se estimará la dificultad de cada ítem con base en las respuestas correctas proporcionadas por alumnos anteriores en la misma prueba.

Punto de partida: Se realizará una estimación inicial del nivel de capacidad que tiene el sujeto a evaluar, para esto, se presentará una pregunta de dificultad media.

Algoritmo de selección de elementos o ítems: Se obtendrán reglas de asociación secuenciales aplicando los algoritmos provistos por WEKA a los conjuntos de datos de los exámenes respondidos por estudiantes de grupos anteriores. De acuerdo con las secuencias de respuestas de los estudiantes actuales, se presentarán las preguntas con una mayor probabilidad de contestar correctamente con base en las reglas de asociación secuenciales con mayor soporte.

Procedimiento de puntuación: Este componente permite obtener la función de probabilidad de la capacidad del sujeto evaluado, mediante las técnicas de estimación de máxima verosimilitud y estimación bayesiana. El método o procedimiento de puntuación es importante, ya que a partir de este se crean intervalos de confianza que sirven para delimitar la precisión del logro.

Criterio de terminación: Uno de los componentes más importantes de un CAT es el criterio de terminación, ya que una prueba o examen adaptativo computarizado solo continuará según el criterio calibrado. En el sistema CAT propuesto, el examen terminará cuando se presenten las preguntas establecidas por el profesor.

3.2. Arquitectura

En la Figura 11 se observa la arquitectura preliminar para el sistema CAT, la cual se basa en el patrón arquitectónico MVC (Modelo-Vista-Controlador). El modelo contendrá las

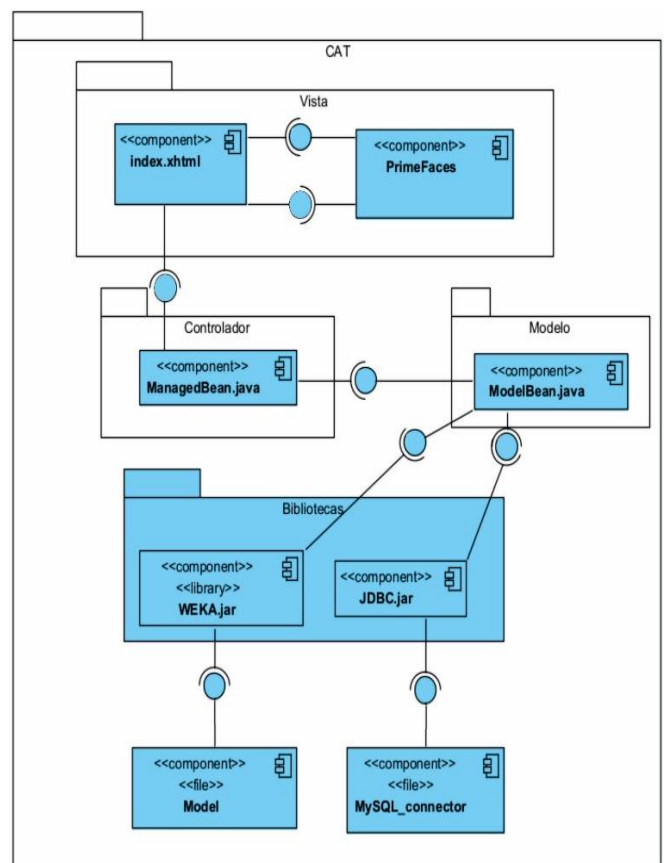


Figura 11: Arquitectura Preliminar del sistema CAT.

clases de Java necesarias para la gestión de los exámenes y de los usuarios. Este se conectará con la herramienta de minería de datos WEKA para la generación de las reglas de asociación secuenciales y con el controlador de MySQL para la gestión de los datos. La vista se implementará con páginas XHTML y el marco de trabajo PrimeFaces. El controlador utilizará los *ManagedBeans* para atender las peticiones y seleccionar las

clases del modelo y las páginas de la vista adecuadas. A continuación, se justifica la elección de cada tecnología.

JavaServer Faces: es un marco de trabajo que está diseñado para el lenguaje de programación Java y ayuda a simplificar el desarrollo web. Este fue elegido debido a que hace una separación entre la presentación de interfaces de usuario y la lógica de la aplicación (Gómez Fuentes y Cervantes Ojeda, 2017), lo que facilitará el mantenimiento del sistema y el crecimiento de este.

MySQL: es un sistema gestor de bases de datos que fue desarrollado por Oracle Corporation y está escrito en los lenguajes de programación C y C++. Este gestor es distribuido por una licencia GNU (Gnu is Not Unix), y es gratuito, cuenta con una comunidad muy activa en Internet (López Herrera, 2016). Además de esto por su velocidad de respuesta en la ejecución de tracciones garantiza el buen rendimiento del sistema propuesto.

WEKA: es una plataforma de software destinada al aprendizaje automático y a la minería de datos que se desarrolló con el lenguaje Java por la universidad de Waikato en Nueva Zelanda. WEKA es capaz de realizar diferentes técnicas de procesamiento de información como lo son: preparación de datos, clasificación, regresión, agrupamiento y minería de reglas de asociación, también permite la instalación de nuevas técnicas de procesamiento como, por ejemplo, de minería de reglas de asociación secuenciales que son una de las partes centrales del sistema, además de permitir la visualización de los datos (Weka 3 - Data Mining with Open Source Machine Learning Software in Java, s.f.).

NetBeans: es un entorno de desarrollo multiplataforma gratuito que es utilizado para el desarrollo de aplicaciones web, aplicaciones de escritorio y aplicaciones móviles, además este entorno integra diferentes lenguajes de programación como Java, PHP, C y C++, también es capaz de manejar diferentes marcos de trabajo, por ejemplo, Symfony para PHP y JavaServer Faces para Java (González Mendoza, 2015), que es el marco seleccionado para el desarrollo de esta propuesta.

UWE: UWE es el acrónimo en inglés de ingeniería web basada en UML, esta metodología tiene como objetivo utilizar el lenguaje común para definir un mapeo a lo largo de diferentes fases que son captura de requisitos, análisis, diseño e implementación. Esta metodología se usa generalmente para el desarrollo web, debido a esto se seleccionó como óptima para el desarrollo de este sistema, ya que una característica importante es que está dirigida por modelos (UWE - UML-Based Web Engineering, s.f.).

4. Conclusión

Las evaluaciones son una parte importante en el proceso de enseñanza-aprendizaje. Al aplicar una prueba adaptativa computarizada se provee a los docentes un conjunto de información útil acerca del nivel de aprendizaje de cada estudiante, por lo que es posible aplicar estrategias para mejorar la calidad de la enseñanza. Como se mencionó anteriormente, los exámenes tradicionales son considerados la forma más común de medir el nivel de aprendizaje de los y las estudiantes, haciendo la suposición de que todos cuentan con

el mismo nivel y aplicándoles el mismo examen, sin embargo, los CAT se adaptan a cada uno de ellos conforme van respondiendo incisos, esto es posible gracias al algoritmo de selección de ítems que integra el CAT donde se aplicará minería de reglas de asociación secuenciales.

Después del análisis realizado de los artículos relacionados con la aplicación de métodos de minería de patrones secuenciales en la educación, se llegó a la conclusión de que el 47% de los trabajos abordó principalmente el problema de la predicción de comportamiento, en contraste, solo el 5% de los trabajos se enfocó en el tema de evaluación, uno de ellos es Cheng et al. (2021) que realizó el análisis de big data y utilizó la teoría de respuesta al ítem en exámenes de enseñanza del idioma inglés obteniendo como resultados un mejor desempeño por parte de los alumnos, mientras que Liu et al. (2018) también abordó el área de enseñanza del idioma inglés, su principal diferencia es que se enfocó en la clasificación de las preguntas con respecto a su sintaxis usando obtención de información y minería de patrones secuenciales, su principal resultado fue un enfoque híbrido para la extracción de características para el control de calidad en inglés. Además de esto el 65% de los trabajos analizados utilizó datos provenientes de estudiantes de nivel superior, sin embargo, ninguno de los trabajos propone el uso de técnicas de minería de patrones secuenciales como método de selección de ítems en pruebas.

En el trabajo futuro se llevará a cabo la implementación del sistema de evaluación adaptativa computarizada, para esto se utilizará el método de minería de reglas de asociación secuenciales GSP como estrategia de selección de ítems en las pruebas, que incluye WEKA, ya que según el trabajo de Tarus et al. (2018) es de los más ampliamente utilizados, por último, se validará el sistema por medio de casos de estudio considerando la evaluación de estudiantes de las asignaturas de Bases de Datos y Programación Lógica y Funcional en una universidad mexicana y se comparará con otros sistemas CAT.

Agradecimientos

Los autores queremos expresar nuestro más profundo agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACYT) por su generoso apoyo, asimismo, deseamos agradecer al Tecnológico Nacional de México (TecNM) por su apoyo institucional para este trabajo.

Referencias

- Aktas, D. E., & Aktas, M. S. (2021). Sequential Rule Mining on the Student Behavior Data of an E-Learning Platform in the Field of Financial Sciences: Case Study. *3rd International Conference on Electrical, Communication and Computer Engineering, ICECCE 2021, June*, 12–13. <https://doi.org/10.1109/ICECCE52056.2021.9514230>
- Al-Twijri, M. I., Luna, J. M., Herrera, F., & Ventura, S. (2022). Course Recommendation based on Sequences: An Evolutionary Search of Emerging Sequential Patterns. *Cognitive Computation*, *14*(4), 1474–1495. <https://doi.org/10.1007/s12559-022-10015-5>
- Anwar, T., & Uma, V. (2022). CD-SPM: Cross-domain book recommendation using sequential pattern mining and rule mining. *Journal of King Saud University - Computer and Information*

- Sciences*, 34(3), 793–800.
<https://doi.org/10.1016/j.jksuci.2019.01.012>
- Bermudez, R. S., Sison, A. M., & Medina, R. P. (2020). Extended PrefixSpan for Efficient Sequential Pattern Mining in a Game-based Learning Environment. *ACM International Conference Proceeding Series*, 118–122. <https://doi.org/10.1145/3379310.3381044>
- Chen, C. M., & Wang, W. F. (2020). Mining Effective Learning Behaviors in a Web-Based Inquiry Science Environment. *Journal of Science Education and Technology*, 29(4), 519–535.
<https://doi.org/10.1007/s10956-020-09833-9>
- Cheng, S. C., Cheng, Y. P., & Huang, Y. M. (2021). To Implement Computerized Adaptive Testing by Automatically Adjusting Item Difficulty Index on Adaptive English Learning Platform. *Journal of Internet Technology*, 22(7), 1599–1607.
<https://doi.org/10.53106/160792642021122207013>
- Cheng Tan, K., Zantedeschi, D., Kumar, A., & Gaspar, A. (2020). *Genetic Algorithm Cleaning in Sequential Data Mining: Analyzing Solutions to Parsons' Puzzles*. 2330–2333. <https://doi.org/10.1145/3520304>
- Czibula, G., Mihai, A., & Crivei, L. M. (2019). SPRAR: A novel relational association rule mining classification model applied for academic performance prediction. *Procedia Computer Science*, 159, 20–29.
<https://doi.org/10.1016/j.procs.2019.09.156>
- Deeva, G., & De Weerd, J. (2019). Understanding Automated Feedback in Learning Processes by Mining Local Patterns. *Lecture Notes in Business Information Processing*, 342, 56–68.
https://doi.org/10.1007/978-3-030-11641-5_5
- Doko, E., Bexheti, L. A., Hamiti, M., & Etemi, B. P. (2018). Sequential pattern mining model to identify the most important or difficult learning topics via mobile technologies. *International Journal of Interactive Mobile Technologies*, 12(4), 109–122.
<https://doi.org/10.3991/ijim.v12i4.9223>
- Fatahi, S., Shabanali-Fami, F., & Moradi, H. (2018). An empirical study of using sequential behavior pattern mining approach to predict learning styles. *Education and Information Technologies*, 23(4), 1427–1445.
<https://doi.org/10.1007/s10639-017-9667-1>
- Gómez Fuentes, M. del C., & Cervantes Ojeda, J. (2017). *Introducción a la Programación Web con Java: JSP y Servlets, JavaServer Faces Obra ganadora del Tercer Concurso para la publicación de libros de texto*. www.cua.uam.mx
- González Mendoza, G. (2015). Herramienta de Desarrollo Netbeans. *Universidad del Norte*, 1–5.
- He, Q., Borgonovi, F., & Paccagnella, M. (2021). Leveraging process data to assess adults' problem-solving skills: Using sequence mining to identify behavioral patterns across digital tasks. *Computers and Education*, 166. <https://doi.org/10.1016/j.compedu.2021.104170>
- Istiyono, E., Dwardaru, W. S. B., Setiawan, R., & Megawati, I. (2020). Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use. *European Journal of Educational Research*, 9(1), 91–101. <https://doi.org/10.12973/eu-jer.9.1.91>
- Klašnja-Miličević, A., Vesin, B., & Ivanović, M. (2018). Social tagging strategy for enhancing e-learning experience. *Computers and Education*, 118, 166–181.
<https://doi.org/10.1016/j.compedu.2017.12.002>
- Kong, M., & Pollock, L. (2020, noviembre 19). Semi-Automatically Mining Students' Common Scratch Programming Behaviors. *ACM International Conference Proceeding Series*.
<https://doi.org/10.1145/3428029.3428034>
- Latypova, V. (2022). Work with Free Response Implementation Process Analysis Based on Sequential Pattern Mining in Engineering Education. *2022 6th International Conference on Information Technologies in Engineering Education, Inforino 2022 - Proceedings*. <https://doi.org/10.1109/Inforino53888.2022.9782969>
- Liu, Y., Yi, X., Chen, R., Zhai, Z., & Gu, J. (2018). Feature extraction based on information gain and sequential pattern for English question classification. *IET Software*, 12(6), 520–526.
<https://doi.org/10.1049/iet-sen.2018.0006>
- López Herrera, P. (2016). Comparación del desempeño de los Sistemas Gestores de Bases de Datos MySQL y PostgreSQL. [Universidad Autónoma del Estado de México]. En *Comparación del desempeño de los Sistemas Gestores de Bases de Datos MySQL y PostgreSQL*. <http://ri.uaemex.mx/handle/20.500.11799/62548%0Ahttp://hdl.handle.net/20.500.11799/62548%0Ahttp://ri.uaemex.mx/bitstream/handle/20.500.11799/62548/TesisPatriciaLopezHerrera.pdf?sequence=3>
- Malekian, D., Bailey, J., & Kennedy, G. (2020). Prediction of students' assessment readiness in online learning environments: The sequence matters. *ACM International Conference Proceeding Series*, 382–391.
<https://doi.org/10.1145/3375462.3375468>
- Mudrick, N. V., Azevedo, R., & Taub, M. (2018). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Computers in Human Behavior*, 96, 223–234.
<https://doi.org/10.1016/j.chb.2018.06.028>
- Niemeijer, K., Feskens, R., Krempel, G., Koops, J., & Brinkhuis, M. J. S. (2020). Constructing and predicting school advice for academic achievement: A comparison of item response theory and machine learning techniques. *ACM International Conference Proceeding Series*, 462–471. <https://doi.org/10.1145/3375462.3375486>
- Norm Lien, Y. C., Wu, W. J., & Lu, Y. L. (2020). How Well Do Teachers Predict Students' Actions in Solving an Ill-Defined Problem in STEM Education: A Solution Using Sequential Pattern Mining. *IEEE Access*, 8, 134976–134986.
<https://doi.org/10.1109/ACCESS.2020.3010168>
- Pacheco-Ortiz, J., Rodríguez-Mazahua, L., Mejía-Miranda, J., Machorro-Cano, I., & Juárez-Martínez, U. (2021). Towards Association Rule-Based Item Selection Strategy in Computerized Adaptive Testing. *Studies in Computational Intelligence*, 966(2), 27–54.
https://doi.org/10.1007/978-3-030-71115-3_2
- Pogorskiy, E., & Beckmann, J. F. (2022). Learners' web navigation behaviour beyond learning management systems: A way into addressing procrastination in online learning? *Computers and Education: Artificial Intelligence*, 3.
<https://doi.org/10.1016/j.caeai.2022.100094>
- Qu, S., Li, K., Wu, B., Zhang, S., & Wang, Y. (2019). Predicting student achievement based on temporal learning behavior in MOOCs. *Applied Sciences (Switzerland)*, 9(24).
<https://doi.org/10.3390/app9245539>
- Real, E. M. H., Pimentel, E. P., & Braga, J. C. (2021). Analysis of Learning Behavior in a Programming Course using Process Mining and Sequential Pattern Mining. *Proceedings - Frontiers in Education Conference, FIE, 2021-October*.
<https://doi.org/10.1109/FIE49875.2021.9637146>
- Salmerón, L. (2011). ¿Por qué realizar un examen mejora nuestro aprendizaje? Lecciones científicas y educativas del efecto del test. *Ciencia Cognitiva. Revista Electrónica de Divulgación*, 5(1), 19–21.
- Shih, W. C. (2018). Mining Sequential Patterns to Explore Users' Learning Behavior in a Visual Programming App. *Proceedings - International Computer Software and Applications Conference*, 2, 126–129.
<https://doi.org/10.1109/COMPSAC.2018.10216>
- Song, W., & Ye, W. (2021). Mining Unexpected Sequential Patterns from MOOC Data. *Proceedings - 12th IEEE International Conference on Big Knowledge, ICBK 2021*, 434–439.
<https://doi.org/10.1109/ICKG52313.2021.00064>
- Song, W., Ye, W., & Fournier-Viger, P. (2022). Mining sequential patterns with flexible constraints from MOOC data. *Applied Intelligence*, 52(14), 16458–16474. <https://doi.org/10.1007/s10489-021-03122-7>
- Tarus, J. K., Niu, Z., & Kalui, D. (2018). A hybrid recommender system for e-learning based on context awareness and sequential pattern mining. *Soft Computing*, 22(8), 2449–2461. <https://doi.org/10.1007/s00500-017-2720-6>
- Taub, M., & Azevedo, R. (2018a). How Does Prior Knowledge Influence Eye Fixations and Sequences of Cognitive and Metacognitive SRL Processes during Learning with an Intelligent Tutoring System? *International Journal of Artificial Intelligence in Education*, 29(1), 1–28. <https://doi.org/10.1007/s40593-018-0165-4>
- Taub, M., & Azevedo, R. (2018b). *Using Sequence Mining to Analyze Metacognitive Monitoring and Scientific Inquiry based on Levels of Efficiency and Emotions during Game-Based Learning*.
- Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction*, 54, 93–103.
<https://doi.org/10.1016/j.learninstruc.2017.08.005>
- UWE - UML-based web engineering. (s/f). Recuperado el 2 de marzo de 2023, de <https://uwe.pst.ifi.lmu.de/index.html>
- Wan, S., & Niu, Z. (2020). A hybrid e-learning recommendation approach based on learners' influence propagation. *IEEE Transactions on Knowledge and Data Engineering*, 32(5), 827–840.
<https://doi.org/10.1109/TKDE.2019.2895033>
- Wang, R., & Zaiane, O. R. (2018). Sequence-Based Approaches to Course Recommender Systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and*

- Lecture Notes in Bioinformatics*, 11029 LNCS, 35–50. https://doi.org/10.1007/978-3-319-98809-2_3
- Weiss, D. J., & Kingsbury, G. G. (2016). *Application of Computerized Adaptive Testing to Educational Problems* Author (s): David J. Weiss and G. Gage Kingsbury Source: *Journal of Educational Measurement*, Vol. 21, No. 4, [Application of Computers to] Published by: National Council on Me. 21(4), 361–375.
- Weka 3 - Data Mining with Open Source Machine Learning Software in Java. (s/f). Recuperado el 2 de marzo de 2023, de <https://www.cs.waikato.ac.nz/ml/weka/index.html>
- Wong, J., Khalil, M., Baars, M., de Koning, B. B., & Paas, F. (2019). Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course. *Computers and Education*, 140. <https://doi.org/10.1016/j.compedu.2019.103595>
- Yang, J. (2021). Effective Learning Behavior of Students' Internet Based on Data Mining. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2021*, 847–850. <https://doi.org/10.1109/ICBAIE52039.2021.9390049>
- Yildirim, D., & Usluel, Y. (2022). Interrelated analysis of interaction, sequential patterns and academic achievement in online learning. En *Australasian Journal of Educational Technology* (Vol. 2022, Número 2).
- Zhang, N., Biswas, G., & Hutchins, N. (2022). Measuring and Analyzing Students' Strategic Learning Behaviors in Open-Ended Learning Environments. *International Journal of Artificial Intelligence in Education*, 32(4), 931–970. <https://doi.org/10.1007/s40593-021-00275-x>
- Zheng, J., Xing, W., & Zhu, G. (2019). Examining sequential patterns of self- and socially shared regulation of STEM learning in a CSCL environment. *Computers and Education*, 136, 34–48. <https://doi.org/10.1016/j.compedu.2019.03.005>
- Zhu, G., Xing, W., & Popov, V. (2019). Uncovering the sequential patterns in transformative and non-transformative discourse during collaborative inquiry learning. *Internet and Higher Education*, 41, 51–61. <https://doi.org/10.1016/j.iheduc.2019.02.001>