



XVI
Congreso Nacional de
Investigación Educativa
CNIE-2021

Detección de estilos de aprendizaje y recomendación personalizada de contenido

Mayra Helena Gómez García

Instituto Tecnológico Superior de Teziutlán
M19TE0018@teziutlan.tecnm.mx

Luis Omar Colombo Mendoza

Instituto Tecnológico Superior de Teziutlán
luis.cm@teziutlan.tecnm.mx

Mario Andrés Paredes Valverde

Instituto Tecnológico Superior de Teziutlán
mario.pv@teziutlan.tecnm.mx

Área temática 18. Tecnologías de la información y la comunicación (TIC) en educación.

Línea temática: Educación, TIC y COVID-19.

Tipo de ponencia: Reporte de investigación parcial o final.



Resumen

La pandemia por coronavirus (COVID-19) ha provocado la mayor interrupción en la historia dentro de los sistemas educativos, afectando a casi 1.600 millones de estudiantes en más de 190 países en todos los continentes.

Debido a esto, los gobiernos deben hacer hincapié en la equidad e inclusión de los estudiantes dentro del sistema educativo, en el caso de México se ha implementado estrategias de educación a distancia.

A pesar del auge que tiene, no basta sólo con hacer uso de herramientas tecnológicas y utilizar contenido multimedia, también es necesario conocer el estado cognitivo de los estudiantes.

Por lo que se propone el uso de una plataforma de aprendizaje que ayude en la detección de estilos de aprendizaje, la cual se apoye en el uso de técnicas de Aprendizaje Automático para Minería Web y técnicas de recomendación personalizada, con el fin de que sea posible realizar ajustes a una o más características del entorno de aprendizaje, por ejemplo, cambios en el contenido de aprendizaje, incorporación de texto, gráficos o videos, etc.

En este sentido la minería web hace referencia al proceso global de descubrimiento de información o conocimiento útil, mediante la extracción de palabras dentro de un documento clasificándolo de acuerdo con su contenido.

En cuanto a los sistemas de recomendación son una herramienta que permite establecer un conjunto de criterios y valoraciones sobre los datos de los usuarios para poder realizar predicciones sobre recomendaciones de elementos que puedan ser de utilidad o valor para el usuario.

Palabras clave: Estilos de aprendizaje, Ambientes virtuales de aprendizaje, Inteligencia artificial, Minería web, Técnicas de recomendación.

Introducción

La pandemia del COVID-19 ha causado el mayor cambio por el que ha pasado la educación a nivel mundial en la historia de la humanidad. Al igual que la economía, la educación se ha visto impactada de manera desfavorable; ante tal situación, los gobiernos del mundo se han visto en la necesidad de cerrar escuelas para tratar de frenar los contagios.

Es por ello por lo que en México se han implementado nuevas formas de trabajo a distancia para que los estudiantes de educación básica continúen con sus procesos de aprendizaje y adquieran los aprendizajes esperados. Una de ellas es el programa público llamado “Aprende en Casa”, con el cual se pretende hacer llegar contenidos educativos a los estudiantes por medio de la televisión, y que se suele acompañar de actividades o sesiones grupales en diversas plataformas digitales educativas y de videoconferencias como Google Classroom, Zoom o Google Meet, de llamadas, videollamadas o mensajes de texto inclusive personales a través de aplicaciones de mensajería instantánea como WhatsApp, así como de cuadernillos de trabajo impresos y/o digitales.

A pesar del auge que en estos momentos tiene esta forma de enseñanza, no basta sólo con que el docente haga uso de herramientas tecnológicas y haga llegar contenido multimedia al estudiante, ya que esto no garantiza un aprendizaje más rápido ni más eficaz; esto depende en gran medida del estilo cognitivo de los estudiantes, debido a que no todos aprenden al mismo ritmo (estilos de aprendizaje), lo cual implica que la estrategia de educación se vea limitada.

Por lo que en este trabajo se propone el desarrollo de una *plataforma digital basada en técnicas de Aprendizaje Automático y de recomendación personalizada que permita evaluar las aptitudes, actitudes y conocimientos de los estudiantes y predecir sus estilos de aprendizaje para crear bancos de contenidos educativos acordes a sus necesidades particulares.*

Conociendo el estilo de aprendizaje de cada alumno será posible ofrecerle contenidos que se adapten a su nivel de conocimiento, incluyendo materiales y recursos que le ayuden a fortalecer sus habilidades y aptitudes y poder alcanzar los aprendizajes esperados de su grado escolar, además de que, con la herramienta aquí propuesta se espera contribuir a disminuir el rezago escolar en nivel básico.

Con este trabajo se busca dar respuesta a la siguiente pregunta de investigación: *¿Es posible sistematizar el proceso de evaluación de las aptitudes, actitudes y conocimientos de los estudiantes para predecir sus estilos de aprendizaje y así recomendar contenidos educativos personalizados de forma automática?*

La respuesta tentativa a esta pregunta está representada por la siguiente hipótesis de trabajo: *‘Con el uso de técnicas de recomendación basada en contenido y de Aprendizaje Automático para minería Web es posible implementar un ambiente virtual adaptativo de aprendizaje que permita sistematizar el proceso de evaluación de aptitudes, actitudes y conocimientos de los estudiantes para recomendar contenido educativo personalizado de forma automática a partir de la predicción de estilos de aprendizaje’.*

A diferencia de los trabajos de vanguardia existentes en la literatura relacionada, en este trabajo pretendemos implementar un enfoque de recomendación simple pero potente: recomendación basada en contenido, que otorgue a los estudiantes la capacidad de describir con sus propias palabras el contenido de los recursos de aprendizaje integrados en la plataforma bajo un enfoque de etiquetado social. De hecho, la mayoría de los trabajos referidos emplean, o bien un enfoque híbrido de recomendación basada en contenido y en Filtrado Colaborativo (Trusthi & Nurjanah, 2017; Nafea et al., 2019), o bien un enfoque de Filtrado Colaborativo puro (Klašnja-Milićević et al., 2011; Bourkhoukou et al., 2016). En el primer caso, los perfiles de elementos de los recursos de aprendizaje, elemento crucial en el enfoque de recomendación basada en contenido, se construyen a partir de “palabras clave” obtenidas mediante cuestionarios aplicados a los estudiantes (Trusthi & Nurjanah, 2017) o en términos de palabras que representan los estilos de aprendizaje considerados por un modelo de estilos de aprendizaje en particular, las cuales son asignadas por un experto del dominio (docente o profesional de la educación) (Nafea et al., 2019).

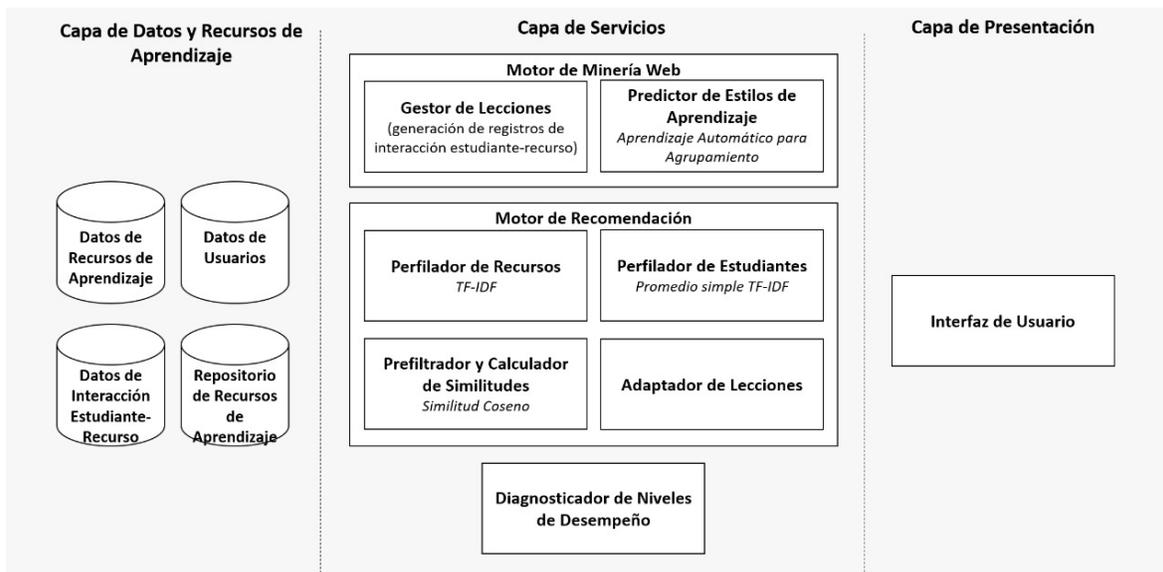
Finalmente, en este trabajo se pretende adaptar los perfiles de usuario de los estudiantes a las características particulares de los estudiantes de educación básica mexicanos, así como a las características propias del contexto de educación a distancia actualmente impuesto en el país por el confinamiento derivado de la pandemia de COVID-19. De ahí que se considere, al menos en principio, el uso del modelo de estilos de aprendizaje VAK (Visual, Auditivo, Kinestésico), modelo referente a nivel de educación básica en México (Trujillo Arteaga, 2017), así como la integración de recursos digitales de aprendizaje provistos actualmente por la Secretaría de Educación Pública del gobierno de México, recursos que han sido verificados y validados por expertos del dominio y que son de acceso completamente abierto y gratuito.

El presente trabajo consta de tres apartados, el primero que es la introducción, en el cual se adentra a la explicación de los fines y las razones que motivaron a dicho trabajo, explicando a detalle lo que se pretende lograr con dicho trabajo. En el segundo apartado que es el de desarrollo, se describe a grandes rasgos la arquitectura de software del ambiente virtual adaptativo de aprendizaje, en el cual se describe a detalle cada una de las capas que conforman el modelo, mencionando la función que cada uno cumple dentro del sistema, así como las herramientas que se utilizaran para la implementación de estos. De igual manera dentro de este mismo apartado se muestran los avances que se tienen hasta el momento, mostrando partes de códigos de la implementación, así como pruebas realizadas. Por último, en el apartado de conclusiones se encuentra una pequeña reflexión en cuanto a los avances alcanzados hasta el momento respecto a los objetivos que persigue el proyecto, y en cuanto a la relevancia social que este pretende lograr.

Desarrollo

A fin de dar cumplimiento al objetivo mencionado en la sección previa, se ha propuesto el diseño de una arquitectura de software de un ambiente virtual adaptativo de aprendizaje. Este diseño, que se puede observar en la Figura 1, se basa en un estilo arquitectónico en capas, con el que se han agrupado de manera lógica los distintos componentes de la arquitectura según sus responsabilidades y las diferentes interrelaciones entre ellos.

Figura 1. Arquitectura de software del ambiente virtual adaptativo de aprendizaje



A continuación, se explican brevemente las funcionalidades de los componentes principales que conforman la arquitectura.

- **Gestor de Lecciones:** este componente es el encargado de permitir al docente la creación y edición de lecciones de contenidos educativos en la plataforma. Ello consiste, prácticamente, en la elección de la temática a abordar, indicada a distintos niveles de profundidad (tema y subtema de primer nivel), dado que la plataforma deberá ser capaz de adaptar las lecciones a cada estudiante de acuerdo con su estilo de aprendizaje y su particular nivel de desempeño. Asimismo, este componente es responsable de generar los registros de interacción de los estudiantes con los recursos de aprendizaje asignados a las lecciones, tarea crucial para la predicción de los estilos de aprendizaje usando técnicas de Aprendizaje Automático para Minería Web.
- **Perfilador de Recursos:** este componente construye perfiles de elementos para los recursos de aprendizaje disponibles en la plataforma. Esta actividad es de vital importancia para la implementación del enfoque y la técnica de recomendación que soporta nuestro enfoque de adaptación de contenidos educativos: el

enfoque de recomendación basada en contenido y la métrica de similitud coseno, respectivamente, según la propuesta de Cantador y colaboradores (Cantador et al., 2010). Para ello, se ha implementado el esquema de ponderación TF-IDF (ver la Formula 1). En este contexto, los términos están representados por etiquetas asignadas por los estudiantes a los recursos bajo un enfoque de etiquetado social (Lavoué, 2011), mientras que, evidentemente, los documentos están representados por los recursos de aprendizaje descritos en términos de dichas etiquetas.

$$tf - iif_{i_n}(t_l) = tf_{i_n}(t_l) * iif(t_l) \quad (1)$$

Donde:

- $tf_{i_n}(t_l)$ es el número de veces que el recurso i_n ha sido anotado con la etiqueta t_l .
- $iif(t_l) = \log \frac{N}{n_i(t_l)}$; $n_i(t_l) = |\{i_n \in I | i_{n,l} > 0\}|$ es
- Perfilador de Estudiantes: este componente está encargado de construir el perfil de usuario de cada estudiante que hace uso de la plataforma. Ello se realiza con base en los perfiles de elementos construidos por el Perfilador de Recursos, específicamente, a partir del cálculo de los promedios de los pesos TF-IDF calculados para todas las etiquetas asignadas a cada uno de los recursos con los que el usuario en cuestión ha interactuado en el pasado. Además, formalmente, los estilos de aprendizaje de los estudiantes, así como sus niveles de desempeño, se consideran parte de sus perfiles de usuario. En este punto cabe mencionar que, inicialmente, esto es, cuando un estudiante es nuevo en la plataforma, se adolece de información sobre la interacción de este con los recursos de aprendizaje, por lo que la plataforma no tiene la habilidad de predecir el estilo de aprendizaje del estudiante en dicha situación. Para resolver este problema se ha optado por utilizar un cuestionario inicial de identificación de estilos de aprendizaje, misma estrategia utilizada para determinar los niveles de desempeño.
- Perfilador y Calculador de Similitudes: como se adelantó previamente, este componente calcula las similitudes entre los perfiles de los estudiantes y los perfiles de los recursos de aprendizaje empleando la métrica de similitud coseno (ver Formula 2). La recomendación de los recursos de aprendizaje más susceptibles de ser del interés de un estudiante en términos de sus preferencias personales, su estilo de aprendizaje y su nivel de desempeño se reduce entonces al cálculo de las similitudes entre el perfil de dicho estudiante y los perfiles de los recursos que son, preferiblemente, hasta el momento desconocidos para él y que se asocian al correspondiente estilo de aprendizaje, así como al correspondiente nivel de desempeño. Esto último supone la realización de un pre-filtrado de recursos basado en estilos de aprendizaje y un pre-filtrado de recursos basado en niveles de desempeño, técnica fundamental en la teoría de los sistemas de recomendación sensibles al contexto (Adomavicius & Tuzhilin, 2015).

$$g(u_m i_n) = \cos_{tf-idf}(u_m i_n) = \frac{\sum_1 tf_{u_m}(t_l) \cdot iuf(t_l) \cdot tf_{i_n}(t_l) \cdot iif(t_l)}{\sqrt{\sum_1 (tf_{u_m}(t_l) \cdot iuf(t_l))^2} \cdot \sqrt{\sum_1 (tf_{i_n}(t_l) \cdot iif(t_l))^2}}$$

Donde:

- $tf_{u_m}(t_l)$ es el número de veces que el usuario u_m ha interactuado con recursos etiquetados con la etiqueta t_l ,
- $iuf(t_l) = \log \frac{M}{n_u(t_l)}$; $n_u(t_l) = |\{u_m \in U | u_{m,l} > 0\}|$ es
- (ver la descripción del componente Perfilador de Recursos).
- **Predicador de Estilos de Aprendizaje:** este componente tiene la responsabilidad de predecir los estilos de aprendizaje de los estudiantes a partir de los registros de interacción de estos con los recursos de aprendizaje empleando técnicas de Aprendizaje Automático. En particular, actualmente se está evaluando la efectividad de distintas técnicas de agrupamiento en la realización de esta tarea. Dichas técnicas de Aprendizaje Automático corresponden al enfoque de Aprendizaje Automático no supervisado, que no supone la necesidad de emplear un conjunto de datos de entrenamiento manualmente etiquetado. Como es posible deducir, la limpieza, tratamiento y reducción (preprocesamiento) de los datos de uso de los recursos de aprendizaje representa un prerrequisito para el uso de las técnicas de Aprendizaje Automático en este contexto. Por último, cabe mencionar que en este trabajo se ha considerado, al menos inicialmente, el modelo de estilos de aprendizaje VAK (Visual, Auditivo, Kinestésico), modelo ampliamente utilizado a nivel de educación básica, y que se ha propuesto el siguiente mapeo entre los estilos de aprendizaje recogidos por dicho modelo y los recursos de aprendizaje a considerar en la plataforma (Bourkougou et al., 2016; El Aissaoui et al., 2018; Azzi et al., 2020) (ver Tabla 1).

Tabla 1. Mapeo de estilos de aprendizaje según el modelo VAK y recursos de aprendizaje

Estilo de Aprendizaje/Categoría de Recurso	Video	Imagen (infografía, mapa, etc.)	Audio	Texto (lectura, glosario, etc.)	Manual/ tutorial
Visual	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Auditivo	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		
Kinestésico	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>

- **Adaptador de Lecciones:** como su nombre lo indica, este componente adaptará los contenidos educativos de las lecciones a las necesidades de cada estudiante con base en las recomendaciones de recursos de

aprendizaje producidas por el Perfilador y Calculador de Similitudes. Como es evidente, estas adaptaciones se reflejarán a nivel de la interfaz de usuario del entorno virtual adaptativo de aprendizaje.

Resultados parciales

Hasta el momento se realizaron pruebas del componente de Perfilador de Recursos, para ello se hizo uso del algoritmo TF-IDF, para su implementación se utilizó el lenguaje de programación JAVA.

Dicha implementación se dividió en dos partes en la primera se realizó el cálculo del TF, en esta parte se determinó la frecuencia relativa de un término específico, el cual puede ser una palabra o una combinación de palabras dentro de un documento. El valor obtenido fue comparado con la frecuencia relativa de todos los demás términos de un texto o documento. El TF contempló la proporción de todas las palabras usadas en el texto.

Dentro de las pruebas se hizo uso de un listado de “palabras clave”, las cuales fueron tomadas como referencia.

```
1 public double tf(List<String> doc, String term) {  
2     double result = 0;  
3     for (String word : doc) {  
4         if (term.equalsIgnoreCase(word))  
5             result++;  
6     }  
7     return result / doc.size();  
8 }
```

La segunda parte de la implementación consistió en el cálculo del IDF, el cual permite obtener la Frecuencia Inversa de Documento, con este fue posible calcular la frecuencia del documento en cuanto a términos específicos, es decir, se determinó la relevancia del documento con respecto a una palabra dentro del listado de palabras específicas.

```
1 public double idf(List<List<String>> docs, String term) {  
2     double n = 0;  
3     for (List<String> doc : docs) {  
4         for (String word : doc) {  
5             if (term.equalsIgnoreCase(word)) {  
6                 n++;  
7                 break;  
8             }  
9         }  
10    }  
11    return Math.log(docs.size() / n);  
12 }
```

Una vez obtenidos los valores de TF-IDF, se realizó el cálculo de estos y se obtuvo la frecuencia de uso para cada palabra clave del listado dentro de los documentos especificados.

```
1 public double tfIdf(List<String> doc, List<List<String>> docs, String term) {  
2     return tf(doc, term) * idf(docs, term);  
3 }
```

De igual manera se trabajó con el componente Perfilador y Calculador de Similitudes, cuya implementación se muestra a continuación:

1	<code>public static double cosineSimilarity(double[] vectorA, double[] vectorB) {</code>
2	<code> double dotProduct = 0.0;</code>
3	<code> double normA = 0.0;</code>
4	<code> double normB = 0.0;</code>
5	<code> for (int i = 0; i < vectorA.length; i++)</code>
6	<code> {</code>
7	<code> dotProduct += vectorA[i] * vectorB[i]; normA += Math.pow(vectorA[i], 2);</code>
8	<code> normB += Math.pow(vectorB[i], 2);</code>
9	<code> }</code>
10	<code> return dotProduct / (Math.sqrt(normA) * Math.sqrt(normB));</code>
11	<code>}</code>

Para el cálculo de la similitud coseno, se utilizaron las palabras claves definidas en dos artículos (vectores), después de ello se tomaron varias de las palabras clave de cada documento, dichas palabras se agruparon en un conjunto, se retomaron los valores de las frecuencias de cada palabra para las palabras contempladas en el conjunto, y se observó que entre mayor es el valor coseno obtenido para el artículo mayor es su similitud, indicándonos en este caso que es el recurso más apropiado para el estudiante.

Conclusiones

Actualmente, la nueva forma de enseñanza-aprendizaje es, más que nunca, una temática estudiada; esto se debe a que los sistemas de enseñanza de todo el mundo se han visto afectados por la pandemia que aqueja al mundo entero (COVID-19), dicha situación está obligando a una reestructuración dentro de los sistemas educativos.

Dicho impacto se está reflejando en el uso de las nuevas tecnologías de la información y comunicación, siendo estas las principales herramientas para hacer llegar la educación a los hogares y con esto lograr que día tras día sean menos estudiantes los que no tengan acceso a la educación.

Como ya se mencionó en países como México se están implementando estrategias para acercar a más estudiantes a los recursos educativos, pero hace falta contar con más herramientas que permitan no sólo hacer llegar dichos recursos a los alumnos, sino ofrecer materiales que se adapten a las necesidades, intereses y niveles cognitivos con los que cuenta el estudiante en curso, es por ello, que el trabajo presentado busca dar

un aporte al campo de la educación ofreciendo las metodologías así como recursos tecnológicos que permitan cubrir ese aspecto dentro de dicho campo,

Las pruebas realizadas hasta el momento demuestran que es posible predecir la relevancia de recursos de aprendizaje utilizados por el estudiante, además de que con esa información se puede generar un perfil del recurso y ubicarlo en cierto campo de conocimiento, área o materia según su relevancia, también es posible generar un perfil del estudiante tomando como base el historial de los recursos a los que este accede, esto con el fin de asegurarse de que cada material proporcionado cubrirá las necesidades de cada estudiante sin importar el nivel cognitivo con el que cuente.

Referencias

- Adomavicius, G., & Tuzhilin, A. (2015). Context-Aware Recommender Systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 191–226). Springer US. https://doi.org/10.1007/978-1-4899-7637-6_6.
- Azzi, I., Jeghal, A., Radouane, A., Yahyaouy, A., & Tairi, H. (2020). A robust classification to predict learning styles in adaptive E-learning systems. *Education and Information Technologies*, 25(1), 437–448. <https://doi.org/10.1007/s10639-019-09956-6>.
- Bourkhouk, O., Bachari, E. E., & Adnani, M. E. (2016). A Personalized E-Learning Based on Recommender System. *International Journal of Learning and Teaching*. <https://doi.org/10.18178/ijlt.2.2.99-103>.
- Cantador, I., Bellogín, A., & Vallet, D. (2010). Content-based recommendation in social tagging systems. *Proceedings of the Fourth ACM Conference on Recommender Systems*, 237–240. <https://doi.org/10.1145/1864708.1864756>.
- El Aissaoui, O., El Madani El Alami, Y., Oughdir, L., & El Alloui, Y. (2018). Integrating web usage mining for an automatic learner profile detection: A learning styles-based approach. *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 1–6. <https://doi.org/10.1109/ISACV.2018.8354021>.
- Klašnja-Miličević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education*, 56(3), 885–899. <https://doi.org/10.1016/j.compedu.2010.11.001>.
- Lavoué, É. (2011). Social Tagging to Enhance Collaborative Learning. In H. Leung, E. Popescu, Y. Cao, R. W. H. Lau, & W. Nejdl (Eds.), *Advances in Web-Based Learning—ICWL 2011* (pp. 92–101). Springer. https://doi.org/10.1007/978-3-642-25813-8_10.
- Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-based Recommender Systems: State of the Art and Trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 73–105). Springer US. https://doi.org/10.1007/978-0-387-85820-3_3.
- Nafea, S. M., Siewe, F., & He, Y. (2019). On Recommendation of Learning Objects Using Felder-Silverman Learning Style Model. *IEEE Access*, 7, 163034–163048. <https://doi.org/10.1109/ACCESS.2019.2935417>.
- Trujillo Arteaga, J. (2017). Estrategias de Enseñanza para implementar según estilos de aprendizaje de los alumnos [Tesis de maestría, Tecnológico de Monterrey]. <https://repositorio.tec.mx/handle/11285/632880>.
- Trusthi, S. L., & Nurjanah, D. (2017). Combination of hybrid filtering and learning style for learning material recommendation. *2017 IEEE Conference on E-Learning, e-Management and e-Services (IC3e)*, 24–29. <https://doi.org/10.1109/IC3e.2017.8409233>.