

Hacia un modelo de evaluación adaptativa personalizada basado en ontologías, contexto y filtrado colaborativo

Oscar Mauricio Salazar Ospina, Demetrio Arturo Ovalle Carranza

Departamento de Ciencias de la Computación y la Decisión
Universidad Nacional de Colombia – Sede Medellín
Medellín, Colombia
{omsalazaro, dovalle}@unal.edu.co

Fernando de la Prieta Pintado

Departamento de Ciencias de la Computación
Universidad de Salamanca
Salamanca, España
fer@usal.es

Resumen— La fase de evaluación desempeña un rol muy importante durante el proceso de enseñanza-aprendizaje de los estudiantes ya que a partir de esta se validan los conocimientos adquiridos por ellos y se descubren falencias y/o fortalezas. Sin embargo, la selección de preguntas por parte del profesor o de la plataforma de aprendizaje no responde a las necesidades, limitaciones y/o características de los estudiantes. En este contexto, la incorporación de mecanismos que permita abstraer de mejor manera las características del usuario para su posterior uso durante el proceso de selección de preguntas trae consigo beneficios como una mejor medición de los conocimientos, un incremento en el interés de los estudiantes, una mejor detección de falencias para la recomendación de nuevos recursos, entre otros. Con el objetivo de realizar una selección de preguntas que responda de mejor manera a las necesidades de los estudiantes, este artículo realiza una caracterización de las técnicas y modelos más relevantes para la selección de preguntas. De igual manera, se propone un modelo ontológico de evaluación adaptativa personalizado apoyado en técnicas de Inteligencia Artificial que incorpora información cognitiva y contextual relevante del estudiante para realizar una mejor selección y clasificación de preguntas durante el proceso de evaluación virtual.

Palabras clave—*evaluación virtual adaptativa; e-assessment, selección y clasificación automática de preguntas, contexto, filtrado colaborativo, perfiles de usuario.*

I. INTRODUCCIÓN

El proceso de evaluación juega un papel fundamental dentro del proceso de aprendizaje del estudiante, ya que permite diagnosticar el estado previo de los conocimientos, monitorear el avance y validar el nivel de conocimiento obtenido al final de la actividad de aprendizaje [1]. En entornos virtuales de aprendizaje este concepto no pierde relevancia y actualmente es objeto de estudio de gran número de trabajos. Dichos esfuerzos han explorado diversas maneras de integrar la evaluación con entornos virtuales y a partir del uso de la tecnología potenciar el proceso, algunos ejemplos son la generación de calificaciones automáticas, la diversificación de tipologías de preguntas, la generación de estadísticas de seguimiento, la retroalimentación continua, entre otros. Sin embargo, la mayoría de entornos de aprendizaje actuales se

centran en modelos de evaluación genéricos que no tienen en cuenta las características de los estudiantes, es decir, todos los estudiantes son evaluados con el mismo esquema sin contemplar gustos, limitaciones, estilos de aprendizaje, niveles cognitivos, contextos sociales, contextos espaciales, etc [2][3][4].

La mayoría de esquemas de evaluación actuales se centran en construir grandes bancos de preguntas que posteriormente son seleccionadas aleatoriamente a los estudiantes, sin considerar, el nivel de dificultad de las mismas. Wauters et al. [5] sostienen que los estudiantes aprenden mejor y se sienten más motivados cuando los ejercicios son seleccionados de acuerdo al nivel de dificultad; ejercicios muy sencillos o sumamente complicados generan desinterés y frustración en los estudiantes. Desafortunadamente, medir el nivel de dificultad de las preguntas a seleccionar dentro de una evaluación no es tarea fácil, debido a que es un ejercicio subjetivo y sumamente personal [6]. Algunos trabajos como el de Watering and Rijt [7] sostienen que los profesores tienden a sobreestimar la dificultad de las preguntas fáciles y subestimar la dificultad de aquellas preguntas difíciles. En contraste, Impara y Plake [8] argumentan que los profesores generalmente clasifican mejor la dificultad de las preguntas difíciles. Algo similar pasa con los estudiantes, algunos consideran fáciles ciertos cuestionamientos, mientras que otros los encuentran difíciles y complicados. Esto corrobora, que el nivel de las preguntas debe estar en función del estudiante y no de la pregunta misma.

Con base en lo anterior, el presente trabajo tiene como objetivo proponer un modelo apoyado en técnicas de Inteligencia Artificial tales como ontologías, lógica difusa y filtrado colaborativo; buscando incorporar información relevante del estudiante que permita realizar una mejor selección de preguntas según el nivel de dificultad durante el proceso de evaluación virtual.

El resto del artículo se organiza de la siguiente manera: en primer lugar, la sección 2 describe los conceptos principales que se encuentran implicados en esta investigación relativa a evaluación adaptativa, ontologías, perfiles de usuario, entre otros. La sección 3 muestra algunos de los trabajos relacionados más representativos en el área de investigación.

En la sección 4 se presenta el modelo de evaluación adaptativa personalizada propuesto. Por último, la sección 5 presenta las conclusiones y el trabajo futuro.

II. MARCO TEÓRICO

A. Evaluación virtual adaptativa

El proceso de evaluación (*assessment*) puede definirse como todas aquellas actividades que los profesores llevan a cabo para ayudar a detectar falencias y limitaciones en los estudiantes en el proceso de enseñanza-aprendizaje y que a su vez permiten cuantificar el progreso del mismo con el fin de retroalimentar el proceso [9]. La evaluación virtual (*e-assessment*) por su parte, es un importante componente del *e-learning* el cual proporciona dirección, enfoque y orientación al alumno mientras participa de algún proceso de aprendizaje virtual [10]. Utilizar este tipo de herramientas automáticas trae ventajas tales como velocidad, disponibilidad, consistencia y objetividad al momento de realizar evaluaciones. Sin embargo, es importante resaltar que estas herramientas requieren de un diseño pedagógico cuidadoso por parte del profesor. Por su parte, la evaluación virtual adaptativa (*adaptive e-assessment*) busca personalizar el proceso de evaluación con el fin de responder a las necesidades de los estudiantes, es por esto que debe estar alineada con el nivel cognitivo, la reputación, las necesidades, las limitaciones, las preferencias y el comportamiento de éstos. El proceso de adaptación comprende dos enfoques, el primero se centra en la ruta de aprendizaje, mientras el segundo se enfoca en el recurso de evaluación mismo [11][12]. El primer enfoque realiza una selección adaptativa de flujos, con el fin de seleccionar las actividades de evaluación que más se adapten al usuario base en experiencias previas o en resultados de evaluación [2]. Por otro lado, el enfoque de personalización de actividades de evaluación se centra en el nodo, es decir, busca la adaptación de la evaluación en términos del nivel cognitivo, reputación, necesidades, limitaciones, preferencias, comportamiento, etc.

B. Perfiles de usuario

Según D'Agostino et al. [13] un perfil es la representación de un objeto particular a partir de las características primordiales que lo describen. En este sentido, podemos definir un perfil de usuario, como la conjunción de datos personales en función del comportamiento, los intereses personales, las limitaciones y las preferencias. Dicho modelo de usuario se caracteriza a su vez por utilizar una estructura de datos adecuada que permita su análisis, recuperación y utilización.

C. Ontologías

Las ontologías buscan obtener representaciones formales de un dominio particular utilizando metodologías bien definidas que permitan abstraer las entidades pertenecientes a dicho dominio, al igual que las relaciones existentes entre ellas [14]. Con base en esto, las ontologías aparecen como mecanismos formales que permiten describir las entidades pertenecientes al dominio de la evaluación virtual, estableciendo relaciones entre los conceptos y permitiendo así, realizar inferencias para la selección y/o clasificación de las preguntas de evaluación según su nivel de dificultad.

III. TRABAJOS RELACIONADOS

A continuación, se presentan algunos trabajos representativos en el área de investigación, estableciendo sus fortalezas y contrastando sus debilidades.

La investigación propuesta por Hajj et al. [15] define un modelo fundamental de *e-assessment* basado en el análisis de las necesidades de los aprendices y de los profesores usando el diagrama de actividades de UML. A partir de este diagrama proponen una aproximación innovadora para especificar un sistema de evaluación personalizada utilizando un workflow. El trabajo también se apoya en ontologías para representar la semántica de las preguntas y así escoger cuestionamientos adecuados.

Una propuesta de sistema de evaluación adaptativa es presentada por Baneres et al. [12], donde las actividades pueden ser adaptadas con base en evidencias recolectadas durante el proceso instruccional de un aprendiz. Cuando se habla de evidencias se hace referencia a puntajes previos, número de ejercicios realizados, detección del plagio e interacciones en foros. A partir de la arquitectura planteada, se realizó una simulación donde se pudo observar el impacto positivo que tiene la selección de actividades de evaluación de acuerdo a las características de los estudiantes.

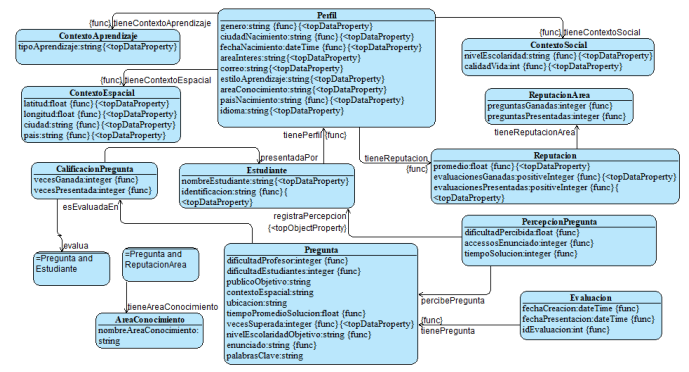


Fig. 1. Taxonomía de la ontología.

Pérez et al. [16] analizan la capacidad de los profesores para determinar el nivel de dificultad de las preguntas asociadas a una evaluación virtual, en contraste con la percepción de dificultad expresada por los estudiantes. El sistema propuesto tiene la capacidad de capturar el nivel de dificultad de las preguntas al momento del diseño instruccional de las evaluaciones, por otra parte, un sistema experto se encarga de capturar la dificultad percibida por los estudiantes durante la prueba en función del tiempo empleado, la calificación obtenida y el número de accesos o lecturas que realizó antes de enviar la respuesta. El trabajo concluye que los estudiantes perciben las preguntas con un mayor grado de dificultad en contraste con el definido por el profesor inicialmente.

Singhal et al. [6] proponen un modelo de dificultad para la generación de preguntas a través de dominios formales de acuerdo al nivel de dificultad otorgado por el usuario. El modelo se apoya en ordenamiento lexicográfico para comparar la dificultad de preguntas basadas en un orden de factores definido por el usuario y un algoritmo asociativo para manejar

estos factores. El sistema desarrollado con base en el modelo permite generar gran número de preguntas con diferentes grados de dificultad considerando únicamente la naturaleza de la estructura de las preguntas.

Viriyadamrongkij y Senivongse [17] proponen un método para medir el nivel de dificultad de las preguntas realizadas en comunidades de desarrolladores en el lenguaje JavaScript basándose directamente en el contenido de la pregunta. Para esto, utilizan la técnica de jerarquía de conceptos y se apoyan en el vocabulario utilizado para estructurar la pregunta. Sin embargo, concluyen que en algunos casos las preguntas con niveles de dificultad alto no necesariamente son ricas en palabras y son complejas por la aplicación de la lógica misma.

Los trabajos previamente descritos evidencian esfuerzos en la búsqueda de mecanismos que permitan seleccionar de mejor manera las preguntas de las evaluaciones. Sin embargo, la mayoría no considera características del estudiante para seleccionar los niveles de dificultad de las evaluaciones, no establecen una estructura formal para describir las entidades asociadas al proceso de evaluación y no incluyen características contextuales del estudiante.

IV. MODELO PROPUESTO

Esta sección presenta el desarrollo del modelo de evaluación adaptativa personalizado basado en la de medición de la dificultad en preguntas de evaluación en función del perfil del estudiante, para su posterior selección en procesos de evaluación virtual. En primera instancia se define un mecanismo formal basado en el concepto de ontologías, que permite la representación de las entidades pertenecientes al dominio de la evaluación virtual junto con sus relaciones. Este modelo ontológico no solo se centra en describir los atributos las evaluaciones y sus preguntas, sino que también considera aspectos del estudiante. A continuación, se presenta en detalle la estructura interna de la ontología de dominio específico desarrollada.

Durante la primera fase de definición del modelo ontológico, se describieron las entidades ligadas al dominio que se querían representar. El primer elemento a describir hace referencia al concepto de la pregunta, al cual fuimos agregando atributos tradicionales tales como: título, descripción, opciones, creador, etc. Una de las grandes dificultades que se encontraron en los modelos actuales es que tratan de definir la dificultad de la pregunta a partir de los atributos de la misma, este trabajo se centra en describir la dificultad de la pregunta en función de las características del usuario y sus percepciones. Con base en esto, el siguiente concepto a describir fue el estudiante a partir de su perfil; para esto, dicho perfil de estudiante describe tres dimensiones: (1) los datos básicos tales como nombre, documento, correo, género, idiomas, estilo de aprendizaje, entre otros. (2) la reputación, la cual se describe a partir de las evaluaciones presentadas, las evaluaciones ganadas, el promedio, el rendimiento en cada una de las áreas del conocimiento definidas, etc. (3) el contexto del estudiante, que a su vez considera tres sub-dimensiones, un contexto espacial (lugar desde donde se interactúa con la evaluación), un contexto social (calidad de vida, nivel de escolaridad, etc.) y un contexto de aprendizaje (define el tipo de evaluación:

formativa, acumulativa, entrenamiento, etc.). A partir de la definición de la estructura del perfil del usuario, se incorporaron nuevos atributos al concepto de pregunta, que ayudan a mejorar el proceso de selección de las mismas. Algunos de estos atributos fueron: dificultad (profesor y estudiantes), público objetivo, contexto espacial, país, tiempo promedio para su solución, nivel de escolaridad objetivo, entre otros. De igual manera, la ontología captura la percepción de la pregunta en términos de dificultad percibida, lecturas realizadas, y tiempo de solución individual. La figura 1 presenta la taxonomía de la ontología desarrollada, los atributos de los conceptos y sus relaciones.

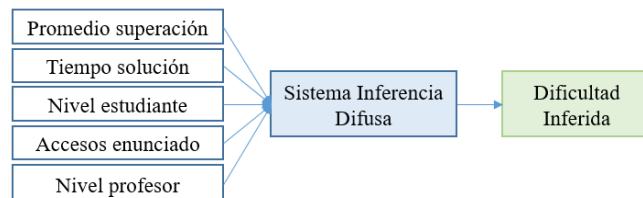


Fig. 2. Módulo de inferencia difusa.

El proceso de categorización de las preguntas a partir del nivel de dificultad se realiza a partir de un sistema de inferencia difusa, que se encarga de recopilar algunos datos consignados en la ontología para procesarlos (ver figura 2). Con base en esto, el sistema construye una bolsa de preguntas inicial con diferentes niveles de dificultad. De igual manera, la ontología consigna la percepción individual del estudiante respecto a cada una de las preguntas que este ha respondido. Apoyándose en esta información, el sistema trata de encontrar similitudes entre perfiles de estudiantes. Este módulo de filtrado colaborativo permite seleccionar preguntas que usuarios similares han calificado previamente, alimentando así la bolsa de preguntas previamente descrita. Posteriormente, un módulo para la categorización de la dificultad en función del perfil de usuario entra en ejecución. Dicho módulo se encarga de definir la dificultad utilizando características que establecen relación entre cada una de las preguntas y el perfil del usuario, estas características son presentadas en la tabla 1.

TABLE I. CARACTERÍSTICAS PARA LA DEFINICIÓN DE LA DIFICULTAD EN FUNCIÓN DEL PERFIL DEL ESTUDIANTE

<i>Perfil de estudiante</i>	<i>Pregunta</i>
Contexto espacial (país)	Ubicación (país), contexto espacial (local, regional, continental, global)
Fecha nacimiento	Público objetivo (niños, adultos, mayores)
Idioma	Idioma
Reputación area del conocimiento	Area del conocimiento
Nivel de escolaridad	Nivel de escolaridad objetivo

Finalmente, un módulo de adaptación de la dificultad se encarga de recibir las respuestas del estudiante y a partir del resultado obtenido, seleccionar la siguiente pregunta. Es decir, el sistema ajusta el nivel a partir de la interacción del usuario con la evaluación. Por otra parte, el sistema cuenta con dos

modalidades, la primera busca ir alternando los niveles de dificultad de acuerdo a las respuestas que el estudiante va enviando. La segunda por su parte, selecciona las preguntas de acuerdo a un nivel de dificultad previamente definido, bien sea por el profesor o por el mismo estudiante. El flujo de procesamiento del sistema es presentado en la figura 3.

V. CONCLUSIONES Y TRABAJO FUTURO

La adaptación de las evaluaciones virtuales en función de las características de los usuarios es un proceso fundamental en la formación debido a que permite identificar de mejor manera las falencias y fortalezas durante el proceso de aprendizaje. Adicionalmente, variar y personalizar el grado de dificultad de las evaluaciones durante su proceso de desarrollo ayuda a mantener el interés de los estudiantes y a hacer su proceso de formación más dinámico y personalizado.

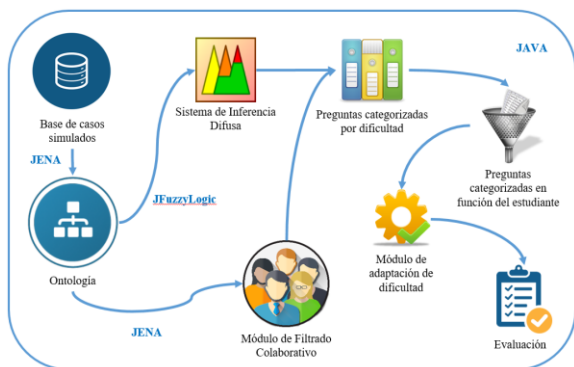


Fig. 3. Flujo de procesamiento del sistema.

El modelo de evaluación adaptativa personalizado propuesto permite controlar el nivel de dificultad en función de características propias del estudiante y del contexto en el cual se desarrolla la actividad de aprendizaje, capturar este tipo de información permite una mejor adaptación de las preguntas y mejora la selección en casos futuros, como por ejemplo la incorporación del módulo de filtrado colaborativo.

Como trabajo futuro se plantea incorporar el modelo propuesto en un sistema de administración de aprendizaje LMS (*learning management system*) que permita recopilar más información de los estudiantes y así poder realizar una mejor validación. De igual manera, se propone utilizar técnicas de aprendizaje de máquina (*machine learning*) para encontrar relaciones en los datos de la ontología, con el fin de encontrar patrones en la información y poder realizar predicciones de mejor manera.

AGRADECIMIENTOS

El trabajo de investigación presentado en este artículo fue financiado parcialmente por la convocatoria 727 de Doctorados Nacionales de 2015 otorgada al estudiante: Oscar Mauricio Salazar Ospina. De igual manera, se agradece especialmente al grupo de investigación BISITE de la Universidad de Salamanca por la beca de pasantía otorgada al estudiante Oscar Mauricio Salazar Ospina.

REFERENCIAS

- [1] D. Monteiro and B. Alturas, "The adoption of e-Recruitment: The Portuguese case: Study of limitations and possibilities by the point of view from candidates and from recruiters," in *Information Systems and Technologies : proceedings of the 7th Iberian Conference on Information Systems and Technologies : (CISTI 2012)*, 2012.
- [2] F. Hajje, Y. B. Hlaoui, and L. J. Ben Ayed, "Personalized and Generic E-assessment Process Based on Cloud Computing," in *2015 IEEE 39th Annual Computer Software and Applications Conference*, 2015, vol. 3, pp. 387–392.
- [3] M. Gusev, S. Ristov, G. Armenski, P. Gushev, and G. Velkoski, "E-Assessment with interactive images," in *2014 IEEE Global Engineering Education Conference (EDUCON)*, 2014, pp. 484–491.
- [4] E. Hettiarachchi, M. A. Huertas, and E. Mor, "E-assessment in high-level cognitive courses: Improving student engagement and results," in *2014 14th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2014, pp. 43–48.
- [5] K. Wauters, P. Desmet, and W. Van den Noortgate, "Adaptive item-based learning environments based on the item response theory: possibilities and challenges," *J. Comput. Assist. Learn.*, vol. 26, no. 6, pp. 549–562, Dec. 2010.
- [6] R. Singhal, S. Goyal, and M. Henz, "User-Defined Difficulty Levels for Automated Question Generation," in *2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)*, 2016, pp. 828–835.
- [7] G. van de Watering and J. van der Rijt, "Teachers' and students' perceptions of assessments: A review and a study into the ability and accuracy of estimating the difficulty levels of assessment items," *Educ. Res. Rev.*, vol. 1, no. 2, pp. 133–147, Jan. 2006.
- [8] J. C. Impara and B. S. Plake, "Teachers' Ability to Estimate Item Difficulty: A Test of the Assumptions in the Angoff Standard Setting Method," *Journal of Educational Measurement*, vol. 35. National Council on Measurement in Education, pp. 69–81, 1998.
- [9] M. Amelung, K. Krieger, and D. Rosner, "E-Assessment as a Service," *IEEE Trans. Learn. Technol.*, vol. 4, no. 2, pp. 162–174, Apr. 2011.
- [10] P. Thomas, D. Haley, A. deRoock, and M. Petre, "e-assessment using latent semantic analysis in the computer science domain: a pilot study," pp. 38–44, Aug. 2004.
- [11] D. Baneres, M. E. Rodríguez, A.-E. Guerrero-Roldán, and X. Baró, "Towards an Adaptive e-Assessment System Based on Trustworthiness," in *Formative Assessment, Learning Data Analytics and Gamification*, Elsevier, 2016, pp. 25–47.
- [12] D. Baneres, X. Baró, A.-E. Guerrero-Roldán, and M. E. Rodríguez, "Towards a General Adaptive e-Assessment System," in *2015 International Conference on Intelligent Networking and Collaborative Systems*, 2015, pp. 314–321.
- [13] E. D'Agostino, A. Casali, R. Corti, and A. Torres, "Sistema de Apoyo al Aprendizaje Diagnóstico Utilizando Perfiles de Usuario : EndoDiag II," in *VIII Simposio Argentino de Informática y Salud*, 2005, pp. 1–14.
- [14] J. Tramullas, A.-I. Sánchez-Casabón, and P. Garrido-Picazo, "An Evaluation based on the Digital Library user: An Experience with Greenstone Software," *Procedia - Soc. Behav. Sci.*, vol. 73, pp. 167–174, Feb. 2013.
- [15] F. Hajje, Y. B. D. Hlaoui, and L. J. Ben Ayed, "A specification approach based on adapted workflow for personalized e-assessment systems," in *Fourth International Conference on Information and Communication Technology and Accessibility (ICTA)*, 2013, pp. 1–6.
- [16] E. V. Perez, L. M. R. Santos, M. J. V. Perez, J. P. de Castro Fernandez, and R. G. Martin, "Automatic classification of question difficulty level: Teachers' estimation vs. students' perception," in *2012 Frontiers in Education Conference Proceedings*, 2012, pp. 1–5.
- [17] N. Viriyadamrongkij and T. Senivongse, "Measuring difficulty levels of JavaScript questions in Question-Answer Community based on concept hierarchy," in *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2017, pp. 1–6.