

DESCUBRIMIENTO DE PATRONES DE NAVEGACIÓN EN OPEN EDX – UNA APROXIMACIÓN ARQUITECTÓNICA

DISCOVERY OF NAVIGATION PATTERNS IN OPEN EDX – AN ARCHITECTURAL APPROACH

Fabián A. Anacona¹, Mario F. Solarte², Gustavo A. Ramírez³

Recibido para publicación: 1 de marzo 2017 - Aceptado para publicación: 9 de mayo 2017

RESUMEN

Los MOOC (Massive Open Online Courses) se han convertido en una estrategia educativa disruptiva en la última década, caracterizada por dar soporte a procesos formativos a gran escala y generalmente de forma gratuita; a pesar de su éxito igualmente tienen debilidades como la rigidez pedagógica pues generalmente las actividades de aprendizaje son las mismas sin importar el perfil y número de los estudiantes. Desde hace algunos años se aplican técnicas de adaptación; pero sin un referente claro sobre cómo realizarla. Este trabajo presenta una aproximación a una arquitectura para descubrir patrones de navegación en los estudiantes en MOOC soportados en la plataforma Open edX, a través del modelo de espacio vectorial y el algoritmo LexRank con Umbral.

Palabras clave: MOOC, SPOC, MPOC, algoritmo, patrones de navegación, Open edX.

ABSTRACT

Massive Open Online Courses (MOOCs) have become a disruptive educational strategy in the last decade, characterized by supporting large-scale and free training processes; in spite of their success they also have debilities such as pedagogical rigidity as generally the learning activities are the same import of the profile and number of the students. Adaptive techniques have been applied for some years; but without a clear reference on how to do it. This work presents an approach to an architecture to discover navigation patterns in students in MOOC supported on the Open edX platform, through the vector space model and the algorithm LexRank with Threshold.

Key words: MOOC, SPOC, MPOC, algorithm, navigation patterns, Open edX.

¹ Ingeniero de Sistemas. Estudiante de Maestría en Ingeniería Telemática. Universidad del Cauca, Popayán, Colombia

² Magister en Ingeniería. Profesor. Universidad del Cauca. Calle 5 # 4-70. Teléfono (+57) 2-809800, extensión 2175. Correo electrónico: msolarte@unicauca.edu.co. Popayán, Colombia.

³ Doctor en Ingeniería Telemática. Profesor. Universidad del Cauca, Popayán, Colombia

1. INTRODUCCIÓN

Los MOOC (Massive Open Online Courses) son una propuesta para universalizar la educación gratuita y de calidad, están basados en el Conectivismo (Siemens 2004). Su principal objetivo es brindar alternativas educativas soportadas en tecnologías de Internet de acceso abierto (Guàrdia et al. 2013, Liyanagunawardena et al. 2013a, Hernández 2013).

Uno de los primeros MOOC ofrecidos fue el de "Conectivismo y conocimiento colectivo" orientado en 2008 por Stephen Downes de la Universidad de Manitoba de Canadá (Mackness et al. 2010); sin embargo, el crecimiento del movimiento MOOC se da con el curso de "Introducción a la Inteligencia Artificial" orientado por Sebastián Thrun, profesor de la Universidad de Stanford de EE.UU y Peter Norvig, Investigador de Google, en el cual se matricularon aproximadamente 160.000 estudiantes (Martin 2012). Los MOOC adquieren fama mundial en el año 2012 de la mano de diversas plataformas como: Udacity, Coursera, edX, entre otras; en las que participan las principales universidades del mundo (Hernández 2013).

Por su parte, edX es una organización de aprendizaje en línea, fundada por la Universidad de Harvard y el Instituto Tecnológico de Massachusetts (MIT) en el año 2012, la cual cuenta con más de 90 socios a nivel mundial, entre entidades sin ánimo de lucro y las principales instituciones que encabezan el ranking de universidades del mundo QS (edX 2013, edX 2016). Open edX es la versión de software libre del sistema de gestión de aprendizaje utilizado en edX.

En Colombia, algunas universidades han realizado MOOC en diferentes plataformas (Semana 2014); entre tanto, la Universidad del Cauca usa Open edX, con nombre de instancia Selene, para ofrecer cursos en línea en ambientes de masividad.

Los MOOC se clasifican de manera general en (Hernández 2013, Guàrdia et al. 2013): cMOOC, cuyo objetivo es la generación de conocimiento de forma colaborativa basándose en la cantidad de aportes generados por sus participantes y los xMOOC, basados en la presentación de contenidos a través de videos cortos, documentos de apoyo, foros de discusión y evaluaciones.

En el año 2013, aparece el término SPOC (Small

Private Online Courses) que son cursos en línea soportados en la plataforma MOOC; pero con la diferencia de no ser abiertos a todo el público y con una cantidad menor de estudiantes (Fox 2013). En el año 2014, aparece el término MPOC (Massive Private Online Courses) cursos en línea igualmente soportados en plataforma MOOC, con la posibilidad de ser masivos, pero sin estar abiertos a todo el público y con una cantidad menor de estudiantes (Guo 2014). Tanto SPOC como MPOC son estrategias encaminadas al reconocimiento académico de las habilidades y conocimientos adquiridos a través de las estrategias de los MOOC.

Los MOOC presentan características similares como: por lo general son de corta duración (Liyanagunawardena et al. 2013b), atraen a miles de estudiantes de diferentes lugares del mundo (Kizilcec 2013), algunos emiten un certificado de aprobación, el cual puede tener un valor monetario (McAuley et al. 2010, Kizilcec 2013).

De igual forma, los MOOC presentan dificultades como: deficiencia pedagógica (Sonwalkar 2013), bajas tasas de retención (Matías González y Pérez Avila 2014) y carencia de atención personalizada (Zapata-Ros 2015); por estas razones Sonwalkar propone el concepto de aMOOC (adaptive MOOC), el cual busca desarrollar un proceso de aprendizaje personalizado permitiendo abrir nuevos campos de investigación, por ejemplo: la identificación de patrones de comportamiento de los estudiantes y la personalización en el desarrollo de un MOOC.

Una alternativa para obtener los patrones de comportamiento es la identificación de las interacciones de los estudiantes con los servicios de la plataforma MOOC, almacenados en el archivo de registro de eventos, por esta razón, en este artículo se presenta una propuesta para encontrar los patrones de navegación en el archivo tracking.log, de la plataforma Open edX a través de técnicas de Minería de Datos Educativos.

Para ello, en la sección 2 se describe la aproximación arquitectónica para el descubrimiento de patrones de navegación, en la sección 3 la estructura del archivo de registro de eventos, en la sección 4 la estrategia del pre-procesamiento requerido, en la sección 5 el algoritmo LexRank con Umbral y en la sección 6 las conclusiones y trabajo futuro.

2. APROXIMACIÓN ARQUITECTÓNICA

La presente propuesta está basada en una arquitectura para identificar perfiles de aprendizaje de los estudiantes, la cual consta de los siguientes componentes: Recopilación de datos, Pre-procesamiento, Extracción de características, Clasificación, Perfil del estudiante, Adaptación (recomendación) presentada por (Hmedna et al. 2017). La presente investigación propone la aproximación arquitectónica mediante la cual es posible descubrir patrones de navegación en cursos masivos soportados en Open edX. Consta de los siguientes módulos: archivo tracking.log, pre procesamiento (archivo de texto plano y representación de estudiantes) y algoritmo LexRank con Umbral; estos se muestran en la Figura 1. Cada uno de los elementos se describe en las siguientes secciones.

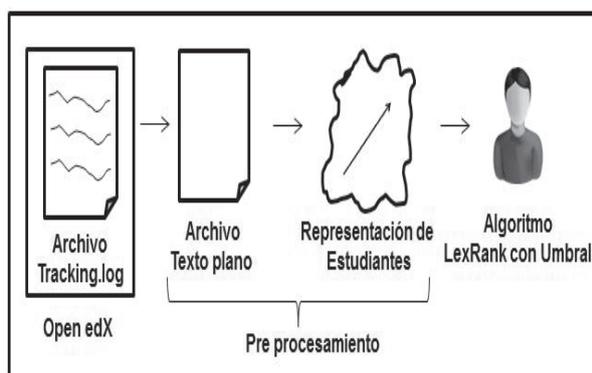


Figura 1. Aproximación arquitectónica para el descubrimiento de patrones de navegación.

2.1. Tracking.log

El archivo Tracking.log se encarga de almacenar todos los eventos realizados sobre la plataforma Open edX en un formato JSON. Los eventos que se almacenan son generados por el servidor, navegador o dispositivo móvil, estos se conocen como emisores. Los eventos se refieren a la información de interacción que se realiza sobre el material del curso y el tablero de instrumentos en el sistema de gestión de aprendizaje de la plataforma.

Para tener acceso a la información se sigue a la ruta /edx/var/log/tracking/, en este lugar se encuentra un conjunto de archivos comprimidos que contienen toda la interacción que los participantes realizan sobre la plataforma. Los eventos generalmente representan una acción

de los emisores sobre la plataforma, a pesar de que dichos eventos pueden representar acciones distintas tienen un conjunto de campos comunes, cuya estructura general se presenta a continuación (Palta y Vásquez, 2016):

```
{
  "username": "", "event_source": "", "name":
  "", "accept_language": "", "time": "", "agent": "",
  "page": "", "host": "", "session": "", "referer": "",
  "context": { "user_id": "", "org_id": "", "course_id":
  "", "path": ""}, "ip": "", "event": "", "event_type":
  "" }
```

Los eventos de los usuarios que permiten conocer cómo fue la navegación de los participantes de un curso sobre una plataforma MOOC, se detallan en las Tablas 1 y 2 (Palta y Vásquez, 2016).

Los eventos seq_goto, seq_next y seq_prev, tienen algunos campos que se describen en la siguiente sección.

Tabla 5. Eventos de navegación en un MOOC. Tomado de (Palta y Vásquez, 2016).

EVENTO	DESCRIPCIÓN
page_close	Este evento se registra desde una acción JavaScript.
seq_goto	Se emite cuando un usuario salta entre las unidades en una secuencia.
seq_next	Se emite cuando un usuario navega a la siguiente unidad en una secuencia.
seq_prev	Se emite cuando un usuario navega a la unidad anterior en una secuencia

Tabla 6. Campos de identificación de pestañas dentro de un MOOC. Tomado de (Palta y Vásquez, 2016).

CAMPO	TIPO	DESCRIPCIÓN
id	Number	Identifica el id edX de la secuencia.
new	Number	Para seq_goto, el índice de la unidad que se saltó al valor especificado. Para seq_next y seq_prev, el índice de la unidad navega al valor especificado.
old	Number	Identifica la pestaña origen del evento.

Dichos eventos están asociados a la navegación de los estudiantes por los distintos servicios de Open edX que pueden depender de la metodología escogida por el diseñador y orientador de un curso, los típicos son ingresos a la plataforma, acceso a videos y otros recursos educativos, participación en foros de discusión (lectura, creación de hilos, respuestas, votación), desarrollo de exámenes y autoevaluaciones, entre otros.

3. PRE-PROCESAMIENTO

El Pre-procesamiento es el primer paso que se realiza en los datos sin procesar, esta etapa tiene como objetivo (Hmedna et al. 2017):

- Limpiar los datos recolectados de información que no aporten.
- Transformar los datos en un formato limpio que pueda ser utilizado por la aplicación.
- Preparar los datos para el análisis por un algoritmo(s), que pertenezca a método de Grafo, Algebraico y Meta Heurístico, entre otros (Anaconda et al. 2015).

En las siguientes sub secciones se describirá cómo transformar el archivo tracking.log a un modelo de espacio vectorial, con lo cual se prepara los datos del archivo para que sean procesados por un algoritmo, como se ilustra en la Figura 2.

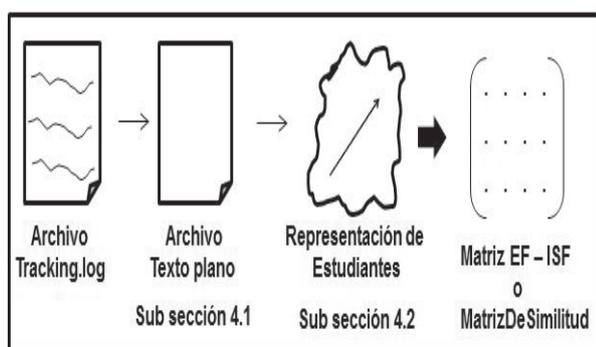


Figura 2. Transformación del archivo Tracking.log en el modelo de espacio vectorial.

3.1. Archivo de texto plano

El archivo Tracking.log se debe leer y procesar para obtener un archivo de texto plano, en este proceso de limpieza y adecuación de los datos se debe eliminar la estructura de formato JSON, manteniendo los eventos de navegación de los estudiantes.

3.2. Representación de Estudiantes

En el documento se propone trabajar con el Modelo de Espacio Vectorial (Salton et al. 1975), porque permite encontrar la relación de los eventos de navegación de los estudiantes en un curso MOOC.

3.2.1. Modelo de Espacio Vectorial

Un vector estudiante \vec{s}_j identifica en qué grado se satisface cada una de las m características del conjunto de estudiantes. Las características (componentes) del vector son un valor numérico. El concepto de característica puede definir la ocurrencia de determinados eventos (Zazo et al. 2002, Manning et al. 2009).

Se considera un espacio vectorial de un estudiante \vec{s}_j , el cual es representado por uno o más eventos e_i ; los eventos pueden ser ponderados (w_{ij}) o no ponderados de acuerdo a su importancia, esto se representa por un m -vector de la siguiente forma (Salton et al. 1975):

$$\vec{s}_{mj} = (w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{mj}) \quad (1)$$

Donde w_{ij} representa el peso del evento i -ésimo del estudiante j .

Cada evento e_i se convierte en una dimensión independiente en el espacio dimensional del vector, la mayoría de los vectores operan en los cuadrantes positivos, dado que a ningún evento se le asigna un valor negativo (Singhal 2001).

3.2.2. Ponderación basada en la frecuencia relativa de un evento

Para la ponderación de los eventos, se usa la Frecuencia Relativa de un Evento (EF-ISF), la cual combina las definiciones de frecuencia del evento ($EF = f_{ij}$) y frecuencia inversa del estudiante ($ISF = \log \frac{N}{n_j}$) para asignar un peso a cada evento del estudiante \vec{s}_j . Se otorga mayor relevancia a los eventos menos frecuentes en el conjunto de estudiantes; pero a su vez más frecuentes en cada estudiante, como se observa en la siguiente ecuación donde w_{ij} representa el peso del evento i -ésimo del estudiante \vec{s}_j (Manning et al. 2009):

$$w_{ij} = (f_{ij}) \left(\log \frac{N}{n_j} \right) \quad (2)$$

3.2.3. Medidas de Similitud

Para el cálculo de similitud entre los estudiantes, estos se deben representar como vectores en

el modelo de espacio vectorial (Alguliev et al. 2011). Teniendo los estudiantes como vectores es posible calcular su semejanza, con medidas de similitud, como la de coseno u otras (Salton et al. 1975). Con similitud de cosenos sería:

Sean \vec{s}_i , \vec{s}_j dos m -vectores diferentes del vector cero, donde el peso de los eventos se calcula con EF-ISF (Ver ecuación (2)). Entonces el ángulo o la similitud entre \vec{s}_i y \vec{s}_j se define en el intervalo $[0, 1]$ de acuerdo a la ecuación (Singhal 2001, Alguliev et al. 2011):

$$\text{simcos}(\vec{s}_i, \vec{s}_j) = \frac{\sum_{k=1}^m w_{ki} w_{kj}}{\sqrt{\sum_{k=1}^m w_{ki}^2 \sum_{k=1}^m w_{kj}^2}} \quad (3)$$

3.2.4. Representación del Traking.log por medio de Matrices

La representación en el espacio multidimensional del conjunto de vectores de estudiantes y la similitud de los mismos, se puede hacer con la Matriz de Eventos por Estudiantes y la Matriz de similitud de Cosenos.

- **Matriz de Eventos por Estudiantes.** La *Matriz EF - ISF* $_{m \times n}$ es una matriz dispersa de pesos dispuestos en m eventos (filas) y n estudiantes (columnas), el elemento ij de la Matriz es denotado por w_{ij} , que corresponde al peso del evento i en el estudiante j , este peso es calculado de acuerdo a la ecuación (2) (Grossman 2001, Gong y Liu 2001).
- **Matriz de Similitud de Cosenos.** La *MatrizDeSimilitud* $_{n \times n}$ de cosenos es una matriz cuadrada, donde n es el número de estudiantes, el elemento ij de la matriz denotado por a_{ij} , es la similitud entre la estudiante \vec{s}_i y \vec{s}_j que aparece en la fila i y la columna j , los elementos de diagonal principal son iguales a 1, ya que es el cálculo de cada estudiante con el mismo, \vec{s}_i con \vec{s}_j cuando $i = j$ (ecuacion (3)) (Grossman 2001).

4. Algoritmo LexRank con Umbral

Al representar los datos del archivo Tracking.log en el modelo de espacio vectorial, estos pueden ser procesados por un algoritmo con algún propósito particular. Se ha considerado el algoritmo LexRank con Umbral para intentar descubrir patrones de comportamiento en la navegación de los estudiantes a través de la arquitectura de información de un curso en Open edX.

Dicho algoritmo pertenece a los Métodos de Grafos, siendo de los más referenciados en generación automática de resúmenes de múltiples documentos; además, al realizar pruebas con los conjuntos de documentos DUC2005¹, DUC2006 y evaluación automática con las medidas ROUGE-1², ROUGE-2 y ROUGE-SU4, con los siguientes algoritmos: Máxima Cobertura Mínima Redundancia (MCMR), método meta heurístico; Análisis Semítico Latente (LSA), método algebraico y LexRank con Umbral, método de grafo; se evidenció que el LexRank con un Umbral de 0.1 obtuvo mejores resultados (Anacona et al. 2015). La Figura 3 muestra el proceso para encontrar el patrón de navegación a través del LexRank con Umbral.

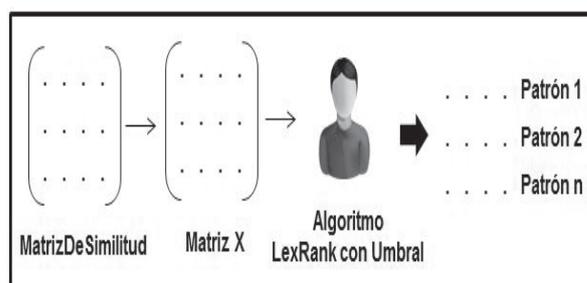


Figura 3. LexRank con Umbral para el cálculo de patrones de navegación.

Erkan y Radev (2004) plantean que un grupo de documentos puede verse como una red de oraciones relacionadas; en la presente investigación se plantea que un estudiante en diferente tiempo con sus respectivos eventos (ver ecuación (1)), puede verse como una red relacionada. Ahora bien, si un estudiante es muy similar a las otras interacciones del mismo estudiante, se puede considerar como el central o sobresaliente.

La similitud de eventos de los estudiantes es representada en una matriz de adyacencia (matriz de similitud de cosenos), ésta se transforma en una matriz X de transición de una cadena de Markov que sea irreducible y aperiódica; luego se encuentra la matriz estacionaria que representa la solución, es decir, los patrones de navegación.

¹ La Conferencia de Entendimiento del Documento (Document Understanding Conference), ofrece un conjunto de documentos para realizar pruebas.

² ROUGE (Recall-Oriented Understudy for Gisting Evaluation), es un estándar para la evaluación automática de resúmenes de textos

A continuación se presenta el pseudo-código del LexRank con Umbral³ para la identificación de patrones de navegación, el Algoritmo 1 resume cómo calcular las puntuaciones LexRank con Umbral para un conjunto de estudiantes.

```

Entrada: Arreglo S de n estudiantes, umbral t, valor de
amortiguamiento dampingFactor
Salida: Arreglo L con las puntuaciones definidos por
LexRank para cada estudiante
Arreglo MatrizDeSimilitud[n][n];
Arreglo L[n];
01 Para i=1 hasta n haga
02     suma=0
03     Para j=1 hasta n haga
04         MatrizDeSimilitud[i][j] = idf-
modificado-coseno(S[i],S[j]);
05         Si MatrizDeSimilitud[i][j] > t haga
06             MatrizDeSimilitud[i][j] =
1;
07             suma++;
08         Si No
09             MatrizDeSimilitud[i][j] =
0;
10         Fin Si
11     Fin Para
15 Fin Para
16 Para i=1 hasta n haga
17     Para j=1 hasta n haga
18         MatrizDeSimilitud[i][j] =
MatrizDeSimilitud[i][j] / suma;
19     Fin Para
20 Fin Para
21 Para i=1 hasta n haga
22     Para j=1 hasta n haga
23         MatrizDeSimilitud[i][j] =
(dampingFactor/n) + (1- dampingFactor)*
MatrizDeSimilitud[i][j];
24     Fin Para
25 Fin Para
26 L = MetodoDePotencia(MatrizDeSimilitud, n, t ); // Ver
Algoritmo 2Método de Potencia.
27 retorne L;

```

Algoritmo 1: Cálculo de puntuación de LexRank con Umbral. Adaptado de (Erkan et al. 2004).

El Método de Potencia describe como se calcula la matriz estacionaria de una Cadena de Markov, mostrado en el Algoritmo 2.

Entrada: Una matriz M estocástica, irreducible y aperiódica

Entrada: Tolerancia de error ϵ

Salida: Vector propio p

01 $p_0 = \frac{1}{N} \mathbf{1}$

02 $t = 0$

03 repita

04 $t = t + 1$

05 $p_t = M^T p_{t-1}$

06 $\delta = ||p_t - p_{t-1}||$

07 hasta que $\delta < \epsilon$;

08 retorne p_t

Algoritmo 2: Método de Potencia. Adaptado de (Erkan et al. 2004).

5. CONCLUSIONES Y TRABAJO FUTURO

Ante la oportunidad de usar la gran cantidad de datos producidos y almacenados en las plataformas de cursos en línea masivos, es posible la aplicación de técnica de Minería de Datos Educativos como del Modelo de Espacio Vectorial combinado con Algoritmos de búsqueda y descubrimiento, encaminados a comprender los hábitos y comportamiento de los estudiantes para mejorar su proceso formativo.

El presente artículo propone una aproximación arquitectónica para descubrir patrones de navegación de los estudiantes en MOOC, MPOC o SPOC alojados en Open edX, a través del estudio de los eventos almacenados en el archivo Traking.log, un pre-procesamiento y posterior uso del algoritmo LexRank con Umbral.

Como trabajo a futuro se plantea la implementación de la arquitectura propuesta y su aplicación a los cursos en línea con reconocimiento académico ofrecidos en la Universidad del Cauca, Popayán, Colombia. Posteriormente se espera poder articular el descubrimiento de patrones de navegación en Open edX con el de los perfiles de los estudiantes (demográficos, estilos de aprendizaje, entre otros) para dar soporte a procesos efectivos de adaptación y personalización del proceso educativo dentro de una plataforma MOOC.

AGRADECIMIENTOS

Los autores agradecen el apoyo recibo por el proyecto MOOC-Maker Construction of Management Capacities of MOOCs in Higher

³ El umbral, clasifica los estudiantes, de los cuales se obtiene el sobresaliente.

Education (561533-EPP-1-2015-1-ESEPPKA2-CBHE-JP) financiado por la Comisión Europea a través del Programa Erasmus+ para la realización y divulgación de los resultados expuestos en el presente artículo.

REFERENCIAS

- [1] G. Siemens, "A learning Theory for the Digital Age," *Elearnspace Everything Learning*, pp. 1-8, 2004.
- [2] L. Guàrdia, M. Maina, and A. Sangrà, "MOOC design principles: A pedagogical approach from the learner's perspective," *eLearning Papers*, 2013.
- [3] T. Liyanagunawardena, S. Williams, and A. Adams, "The impact and reach of MOOCs: a developing countries' perspective," *eLearning Papers*, 2013.
- [4] S. Hernández, "SCOPEO INFORME N°2: MOOC: Estado de la situación actual, posibilidades, retos y futuro," 2013.
- [5] J. Mackness, S. Mak, and R. Williams, "The ideals and reality of participating in a MOOC," 2010.
- [6] F. G. Martin, "Will massive open online courses change how we teach?," *Communications of the ACM*, vol. 55, pp. 26-28, 2012.
- [7] edX. (2013, 24-09-2016). Release of XBlock Source Code. Available: <https://www.edx.org/press/edx-takes-first-step-toward-open-source>
- [8] edX. (2016, 30-10-2016). Quality education for everyone, everywhere Available: <https://www.edx.org/about-us>
- [9] Semana. (2014, 22-02). Las clases que no se quedan sin cupos. Available: <http://www.semana.com/educacion/articulo/cursos-en-linea-masivos-abiertos-en-colombia/409366-3>
- [10] A. Fox, "From moocs to spocs," *Communications of the ACM*, vol. 56, pp. 38-40, 2013.
- [11] W. Guo, "From SPOC to MPOC--The Effective Practice of Peking University Online Teacher Training," in *Educational Innovation through Technology (EITT)*, 2014 International Conference of, 2014, pp. 258-264.
- [12] T. R. Liyanagunawardena, A. A. Adams, and S. A. Williams, "MOOCs: A systematic study of the published literature 2008-2012," *The International Review of Research in Open and Distributed Learning*, vol. 14, pp. 202-227, 2013.
- [13] R. F. Kizilcec, "Collaborative Learning in Geographically Distributed and In-person Groups," in *Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED)*, 2013, pp. 67-74.
- [14] A. McAuley, B. Stewart, G. Siemens, and D. Cormier, "The MOOC model for digital practice," 2010.
- [15] N. Sonwalkar, "The first adaptive MOOC: A case study on pedagogy framework and scalable cloud Architecture—Part I," in *MOOCs Forum*, 2013, pp. 22-29.
- [16] H. Matías González and A. Pérez Avila, "Los Cursos En Línea Masivos Y Abiertos (MOOC) Como Alternativa Para La Educación a Distancia (Massive Open Online Courses (MOOC), an Alternative to Distance Learning)," 2014.
- [17] M. Zapata-Ros, "MOOCs, una visión crítica y una alternativa complementaria: La individualización del aprendizaje y de la ayuda pedagógica," *Campus virtuales*, vol. 2, pp. 20-38, 2015.
- [18] B. Hmedna, A. El Mezouary, O. Baz, and D. Mammass, "Identifying and tracking learning styles in MOOCs: A neural networks approach," *International Journal of Innovation and Applied Studies*, vol. 19, p. 267, 2017.
- [19] R. Palta and J. Vásquez, "Descubrimiento de patrones de interacción en cursos MOOC en entornos ONLINE: Un enfoque utilizando Minería de Procesos. Caso

de Estudio: "Curso de la metodología DICREVOA en Open edX", Facultad de Ingeniería, Universidad de Cuenca, 2016.

- [20] F. A. Anacona Anacona, C. A. Cobos Lozada, and M. E. Mendoza Becerra, "Algoritmo para generación automática de resúmenes extractivos genéricos de múltiples documentos basado en consensos," Monografía, Departamento de Sistemas, Universidad del Cauca, 2015.
- [21] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Commun. ACM*, vol. 18, pp. 613-620, 1975.
- [22] Á. F. Zazo Rodríguez, C. G. Figuerola Paniagua, J. L. A. Berrocal, and R. Gómez Díaz, "Recuperación de Información Utilizando el Modelo Vectorial. Participación en el taller CLEF-2001," Universidad de Salamanca Mayo 2002.
- [23] C. D. Manning, P. Raghavan, and H. Schütze, *An introduction to information retrieval*. Cambridge: Cambridge University Press, 2009.
- [24] A. Singhal, "Modern Information Retrieval: A Brief Overview," *IEEE Data Eng. Bull.*, vol. 24, pp. 35-43, 2001.
- [25] R. M. Alguliev, R. M. Aliguliyev, M. S. Hajirahimova, and C. A. Mehdiyev, "MCMR: Maximum coverage and minimum redundant text summarization model," *Expert Systems with Applications*, vol. 38, pp. 14514-14522, 2011.
- [26] S. I. Grossman, *Álgebra Lineal: Editorial Offset*, S.A. de C.V, Durazno No. 1 esq. Ejido, Col. Las Peritas, Tepepan Xochimilco, C.P. 16010 México D.F, 2001.
- [27] Y. Gong and X. Liu, "Generic text summarization using relevance measure and latent semantic analysis," presented at the Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, New Orleans, Louisiana, USA, 2001.
- [28] G. Erkan and D. R. Radev, "LexRank: graph-based lexical centrality as salience in text summarization," *J. Artif. Int. Res.*, vol. 22, pp. 457-479, 2004.