

**Mecanismo formal de clasificación automática de  
un repositorio de objetos de aprendizaje, a partir de una valoración individual de los objetos, con el fin  
de mantenerlos en un ciclo permanente de mejoramiento  
(Mecanismo de clasificación de objetos de aprendizaje)**

**Formal mechanism for automatic classification  
a repository of learning objects, from an individual assessment of objects in order to keep in a continuous  
cycle of improvement**

Esp. Julián Esteban Gutiérrez Posada<sup>1</sup>, PhD. Miguel Francisco Crespo Alvarado<sup>2</sup>

*En este artículo se desarrolla principalmente un mecanismo formal de clasificación automática de un conjunto de objetos de aprendizaje según el valor de unos indicadores. Esto implica construir un espacio de medida, una función de medida, una función de disimilaridad y un método de clasificación. Adicional a esto, se aplicó un proceso de medición que ayuda a dar mayor sentido a la información obtenida por cada objeto y por las posibles aglomeraciones obtenidas como producto de la clasificación.*

*This article mainly develops a formal mechanism for automatically classifying a set of learning objects according to the value of some indicators. This involves building an area of measurement, a measurement function, a function of dissimilarity and a method of classification. Additional to this, we applied a measurement process that helps to give greater meaning to the information obtained by each object and possible clusters obtained in this classification.*

**Palabras clave: mecanismo de clasificación, objetos de aprendizaje, clasificación automática**

## I. INTRODUCCIÓN

LA manera en la que actuamos en el mundo; esto es, en nuestra vida cotidiana, en las diferentes actividades que abordamos en el día a día; depende directamente de la imagen que sobre la realidad tenemos. Aunque se habla de una distancia entre percepción y realidad, en los hechos, nuestra acción no distingue entre una y otra. Así, percepción y realidad, terminan siendo lo mismo.

Ahora bien, la manera en la que comprendemos el mundo, es decir, nuestra percepción, es determinada por las experiencias que vamos teniendo en la vida y por la manera en que nos apropiamos de ellas. Una muy importante fuente de experiencias es la escuela en todos sus niveles.

La educación, podría entenderse como el proceso mediante el cual los individuos hacen suya una cultura, esto es, se apropian de una forma característica de comprender la realidad y de actuar en ella. Obviamente que la cultura que es transmitida en los procesos educativos no es un ente fijo ni monolítico; sino que hay en ella una gran diversidad. No obstante, nadie puede trascender los límites propios que la cultura le presenta.

Por tanto, cuando nos educamos recibimos una forma posible de vivir nuestra cultura.

En el presente, los procesos educativos formales de transmisión de la cultura se han visto enriquecidos por la aparición de nuevas fuentes informativas que abren, a quienes se están formando, una multiplicidad de opciones culturales. Ello ha implicado la necesidad de transformar las formas en las que se organiza la educación, para dar cabida a esa amplia variedad de formas culturales que son transmitidas, por ejemplo a través de los "objetos de aprendizaje" que circulan en la Internet. Un objeto de aprendizaje es un "recurso digital que puede ser reutilizado para ayudar en el aprendizaje"[44] para formación de sus estudiantes, profesores o administrativos.

---

<sup>1</sup> Profesor de la Universidad de la Quindío en Programa de Ingeniería de Sistemas y Computación y asesor de la Unidad de Virtualización. [jugutier@uniquindio.edu.co](mailto:jgutier@uniquindio.edu.co)

<sup>2</sup> Profesor de la Universidad Autónoma de Bucaramanga (UNAB) en la Maestría de E-Learning. [migrespo@hotmail.com](mailto:migrespo@hotmail.com)

Pero, la transformación organizativa necesaria para la inclusión de esos objetos de aprendizaje, trasciende lo meramente educativo hacia aspectos que conciernen a la administración de las instituciones de enseñanza. Las organizaciones educativas que hacen uso de estos objetos de aprendizaje, como en las universidades, deben mantener un repositorio/banco formal o informal de estos objetos para facilitar su utilización.

Un ejemplo de un repositorio de objetos de aprendizaje es: “El Banco Nacional de Objetos de Aprendizaje del Ministerio de Educación Nacional de Colombia”<sup>3</sup>.

Ahora bien, los objetos de aprendizaje como cualquier otra forma de transmisión cultural, contribuyen, como lo hemos venido afirmando, a crear en las personas una imagen consciente o inconsciente de la realidad que en ellos se trata. Esta imagen puede ser fiel o distorsionada dependiendo de la calidad del mismo; cuando por "calidad", comprendemos el grado de fidelidad con el que la información ofrecida por el objeto de aprendizaje, se apega a la realidad; aún en el entendido de que la "realidad" es algo que tiene una multiplicidad de caras.

El problema al que queremos apuntar es el siguiente: Dado que nuestra acción depende de la manera en la que comprendemos la realidad, ¿qué ocurriría si esa comprensión fuera mayoritariamente distorsionada? En otras palabras, si constantemente nos mantuviéramos recibiendo una información equivocada sobre el acontecer, nuestra acción se correspondería con el error; por tanto, tendería a ser una acción equivocada. En una sociedad en la que todo el mundo o la mayoría actúa de manera equivocada, las consecuencias serían, como lo podemos suponer, desastrosas.

Sería entonces ideal, poder disponer de un mecanismo formal de clasificación de los objetos almacenados en un repositorio, para determinar por ejemplo: ¿Cuáles de ellos deben ser mejorados? ¿Cuáles deben ser descartados? ¿Cuáles son los mejores objetos que se disponen sobre cierto tema? ¿Cuáles no lo son? Sin embargo, una mirada rápida al “Banco Nacional de Objetos de Aprendizaje del Ministerio de Educación de Colombia” no evidencia la existencia de un mecanismo de clasificación. Por supuesto, un mecanismo así no garantizaría directamente la calidad de la información contenida en los objetos, pero seguramente ayudaría a determinar cuáles deben pasar por mecanismos más detallados de evaluación y mejoramiento, y cuales todavía no. En todo caso, podría por lo menos mantener al educando y al educador en un "estado de alerta", que le permitiría no asumir como verdad absoluta aquello que está expuesto en los objetos de aprendizaje no certificados.

Por otra parte, mantener en un ciclo de mejoramiento continuo de los objetos que salieron clasificados en una categoría no muy buena, permitirá seguramente que las personas o usuarios de estos repositorios cuenten con objetos cada vez mejores, disponiendo de muy buenos objetos de aprendizaje lo que ayudaría por ejemplo a:

- Que los usuarios de estos objetos se formen una visión menos distorsionada de la realidad, disminuyendo así el impacto que estos objetos de baja calidad ejercerían sobre el comportamiento individual o colectivo.
- Crear un ambiente de mayor confianza en el uso de los objetos de aprendizaje y en las habilidades o conceptos generados a partir de su utilización.

---

<sup>3</sup> <http://64.76.190.172/drupalM/>

- Focalizar recursos tanto humanos como económicos en mejorar los actuales “peores objetos” del repositorio.

Por lo tanto la pregunta que nos hicimos fue: **¿Es posible definir un mecanismo formal de clasificación automática de un repositorio de objetos de aprendizaje, a partir de una valoración individual de los objetos, con el fin de mantenerlos en un ciclo permanente de mejoramiento?**

Por supuesto, lo ideal de este mecanismo formal de clasificación es que no dependa del significado de los indicadores a utilizar, sino del valor asignado a ellos. Así, una institución podrá utilizar ciertos indicadores y otra por alguna razón, podrá utilizar otros indicadores y el mecanismo de clasificación debe operar indistintamente, por supuesto, todos los objetos a clasificar deben tener evaluados todos los indicadores a utilizar.

## II. TRABAJOS PREVIOS

Muy pocas personas en el mundo podrían pensar que la calidad en general no es importante, más aun, muy pocas personas pensarían que la calidad es algo que se consigue de forma instantánea. Casi de manera generalizada se acepta que la calidad es un proceso continuo, lento y permanente. También se acepta que uno de los principales enemigos de la calidad es la complacencia. Y es ahí en donde tiene lugar este trabajo de investigación de la línea de Pensamiento Sistémico de la Universidad Autónoma de Bucaramanga (UNAB).

En el mundo existen muchas colecciones de objetos de aprendizaje, es más, existen repositorios o bancos de objetos de aprendizaje como el “Banco Nacional de Objetos de Aprendizaje del Ministerio de Educación de Colombia”, en ellos, es común observar mecanismos de búsqueda y seguramente mecanismos de clasificación sobre ciertas temáticas particulares. Sin embargo, una vez un objeto es ingresado al repositorio, esté permanece constante y a la espera de ser consultado, tal y como permanece un libro en una biblioteca.

Un punto a resaltar es que los resultados de los mecanismos de búsqueda pueden ser manipulados, reordenados o priorizados según criterios ajenos a un sistema formal, ya sea por cuestiones políticas, económicas o sociales. Estas posibles manipulaciones han logrado despertar preocupaciones en algunas personas como Micah White (activista independiente que escribe en Adbusters Magazine<sup>4</sup>.) quien piensa que "quién controla los resultados de la búsqueda controla también al autor de la búsqueda", pues visualiza un futuro en donde la diversidad de las ideas desaparezca y "todos" pensemos y en últimas actuemos tal y como quieren los que controlan los motores de búsqueda como google.

Por otro lado, si lo que se desea es tener un repositorio de objetos de aprendizaje de calidad, es necesario entrar en el ciclo continuo de mejoramiento, es decir, buscar los objetos del repositorio que merecen ser mejorados, extraerlos, mejorarlos e ingresarlos nuevamente al repositorio, reclasificándolos y encontrando los nuevos objetos a mejorar, y así sucesivamente... Por otro lado, uno de los grandes enemigos de la calidad es la falta de constancia, pues los movimientos de mejoramiento de calidad son procesos de largo plazo.

Este movimiento hacia la calidad total, no se trata únicamente de mejorar lo que está mal, se trata también de

---

<sup>4</sup> [www.micahmwhite.com](http://www.micahmwhite.com)

encontrar sus causas para así poder eliminar el problema de raíz. Y es en este punto en donde los procesos de clasificación y aglomeración (cluster) tienen verdadera importancia, pues una vez clasificados los objetos de aprendizaje, se pueden identificar no sólo los objetos a mejorar, sino grupos de objetos y determinar las características que son comunes en ellos y así poder tratar de encontrar las posibles causas que provocan esa falta de calidad en el repositorio.

Sin lugar a dudas, existe innumerable bibliografía sobre calidad, calidad total, procedimientos de clasificación y procedimientos de aglomeración, sin embargo no sea podido encontrar bibliografía que combine estas temáticas para resolver el problema ya mencionado. Sí se ha encontrado bibliografía que hable sobre clasificación de documentos, que en últimas tiene algo de similar al contexto general de lo ya expuesto.

En el artículo “Survey of clustering data mining techniques” [2] se puede ver un estudio de las técnicas de agrupación de Data Mining, al profundizar en una clasificación de los algoritmos de aglomeramiento (Cluster) más conocidos, sin intentar ser exhaustiva. Este artículo toma un interés significativo para esta investigación, pues los algoritmos citados, pueden ser utilizados de forma general en distintos ámbitos, como en este caso en la clasificación de objetos de aprendizaje.

Existen sistemas más especializados como por ejemplo para clasificar documentos Web, como lo muestra [16] en su artículo “Efficient phrase-based document indexing for Web document clustering”. En él, los autores presentan un sistema compuesto por cuatro componentes (identificación de la estructura del documento, representación gráfica del índice del documento, cálculo de similaridad entre documentos y cluster incremental) en un intento de mejorar la solución al problema de la agrupación de documentos en un dominio web. Ellos indican que las aplicaciones potenciales incluyen: la agrupación automática de resultados de un motor de búsqueda, la recuperación de la información, la medición de la similitud de documentos, entre muchos otros. Como se puede apreciar entre las aplicaciones se encuentran la similaridad de documentos, estos sistemas y similares, analizan el contenido del documento y usando algún criterio miden su similitud con otros documentos, y así logran aglomerarlos según el grado de similaridad encontrados.

En el artículo “Learning Objects Clustering based on Semantic Understanding of Text” [42], se puede observar otro sistema que mide la similaridad entre elementos, pero esta vez sobre objetos de aprendizaje basados sobre la comprensión de los textos internos al objeto de aprendizaje. Ellos indican que la calidad de la agrupación lograda con este modelo supera el modelo tradicional de espacios vectoriales.

No siempre se necesita medir la similaridad entre documentos, sino también sobre imágenes y este es el caso que se observa en el artículo “Learning object representations by clustering banana wavelet responses” [35], en el cual se introduce un sistema que aprende como representar un objeto a partir de imágenes de entrenamiento. Cada imagen es pre-procesada usando Banana Wavelets, el resultado es una representación del objeto por medio de líneas curvas. Estas líneas se insertan en un algoritmo de clasificación que aprende a separar las características específicas del objeto.

Como se ha podido observar hasta ahora, la clasificación se ha realizado usando como insumo el contenido de los documentos o de las imágenes, sin embargo en muchas ocasiones, se necesitan clasificar elementos por el valor de un conjunto de elementos (frequent itemsets), como por ejemplo el valor obtenido por un objeto de aprendizaje en un conjunto de indicadores.

El artículo “Using Conjunction of Attribute Values for Classification”[11] presenta una serie de algoritmos de clasificación que usan "frequent itemsets". Otro artículo es “Large hierarchical document clustering using frequent itemsets” [4] el cual presenta un nuevo enfoque para abordar la clasificación de documentos con requisitos especiales, tales como: la dimensionalidad, el alto volumen, y la facilidad de la navegación significativa por los clusters. Ellos definen un criterio de agrupación usando algunas palabras comunes.

Se podría continuar así por mucho tiempo, citando uno y otro artículo que mencione algoritmos o criterios diversos de clasificación, para clasificar diversos tipos de objetos, ya sean documentos, imágenes, incluso archivos de sonido. De nuevo entre toda la bibliografía revisada, no se pudo encontrar ninguna que resolviera el problema aquí planteado.

### III. MECANISMO AUTOMÁTICO DE CLASIFICACIÓN

#### A. Espacio de medida

Antes de comenzar a definir formalmente el espacio de medida, es necesario hacer tres definiciones fundamentales basadas en el libro de Bruce K. Driver “Analysis Tools with Applications”[10].

Sea  $X$  un conjunto. Una sigma álgebra de subconjuntos de  $X$  es una clase  $\gamma$  que cumple las siguientes propiedades:

- $X \in \gamma$
- Si  $A \in \gamma$ , entonces  $A^c \in \gamma$
- Si  $A_1, A_2, \dots \in \gamma$ , entonces  $\bigcup_{n=1}^{\infty} A_n \in \gamma$

Una función de medida  $\mu$  definida sobre  $\gamma$  es una aplicación de  $\mu: \gamma \rightarrow [0, +\infty)$   $[0; +1)$  que cumple las siguientes propiedades:

- $\mu(\emptyset) = 0$ ,
- Si  $A_1, A_2, \dots$  son disyuntos entonces

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n), \quad \forall A_i \in \gamma$$

Sea  $X$  un conjunto. Sea  $\sigma$  una sigma álgebra de subconjuntos de  $X$ . Sea  $\mu$  una función de medida en  $\sigma$ . Entonces se define un espacio de medida como la terna  $\langle X, \sigma, \mu \rangle$ .

Una vez hechas estas definiciones fundamentales, se procede a definir, como aporte de esta tesis, el espacio de medida para objetos de aprendizaje de la siguiente forma:

Un objeto de aprendizaje  $O$  se define como la terna  $\langle C, S, P \rangle$  en donde  $C$  es el contenido del objeto de aprendizaje,  $S$  es el conjunto de los valores que toman los atributos definidos en el metadato asociado al objeto y  $P$  es el conjunto de todos los pares ordenados  $(i, v_i)$  en donde  $i$  es el código que identifica al indicador y  $v_i$  es el valor asociado a él.

Un repositorio de objetos de aprendizaje  $R$  se define como  $R = \{O_1, O_1, \dots, O_m\}$ ,  $m \geq 0$  Si  $m = 0$ , se dice que el repositorio se encuentra vacío.

A partir de aquí considere  $X$  como el repositorio  $R$  y el sigma álgebra como partes de  $R$ . Como ya es sabido, partes de  $R$  es la mayor sigma álgebra que puede ser construida a partir de este conjunto.

Sea  $T$  el conjunto de categorías de entidad para objetos de aprendizaje. Entonces  $\forall x \in R$ ,  $x$  debe tener asociada una y sólo una categoría  $t \in T$ .

Algunas categorías de entidad para objetos de aprendizaje pueden ser: presentaciones, videos, unidades, cursos.

Una categoría de entidad  $t$  para un objeto de aprendizaje, se define como la terna  $\langle n, m, V \rangle$  en donde  $n$  es el nombre la categoría,  $m = \|\{x \in R : x \text{ este asociada a la categoría } t\}\|$ , y  $V$  es el conjunto de todas las ternas  $\langle i, d, v \rangle$  en donde  $i$  es el identificador del indicador,  $d$  representa la descripción del identificador y  $v$  es la suma de este indicador  $\forall x \in R$  y que  $x$  este asociado a la categoría  $t$ .

Para este mecanismo de clasificación en particular, se consideran sólo los indicadores que puedan ser valorados numéricamente. Sin embargo esto no debe ser visto como una debilidad, por el contrario, debe ser visto como una oportunidad para futuras investigaciones. La idea, sería crear una nueva función de medida, que por supuesto hay que demostrar matemáticamente, tal y como se hará a continuación con la función propuesta.

Sea la función  $h(x) = \|V^{(x)}\|$  definida como  $h: X \rightarrow Z^+$ , en donde  $\|V^{(x)}\|$  representa la cardinalidad del conjunto  $V$  de la categoría de entidad asociada al objeto  $x$ .

En otras palabras,  $h(x)$  indica el número de indicadores asociados a la categoría de clase a la que pertenece el objeto de aprendizaje  $x$ .

Sea la función  $g$ , definida como  $g: X \rightarrow [0, +\infty)$ , en donde  $g(x)$  representa la cantidad de indicadores del objeto de aprendizaje  $x$ , cuyo valor en  $P$ , se encuentra por encima del promedio (inclusive) de este indicador en la categoría de entidad a la que corresponde el objeto de aprendizaje  $x$ .

Para hacer esto, se debe consultar  $V$  y usando el valor  $m$ , calcula el valor del promedio, en el  $t$  asociado al objeto de aprendizaje  $x$ .

Se define la función de medida  $\mu$  (porcentaje de aceptación) sobre un elemento de  $\gamma$  como una aplicación de

$$\mu: \gamma \rightarrow [0, +\infty), \text{ así: } \mu(Y) = \begin{cases} \sum_{n=1}^{\|Y\|} \frac{g(y_n)}{h(y_n)}, & \text{Si } \|Y\| \geq 1, \forall Y \in \gamma \text{ y } y_i \in Y \\ 0, & \text{Si } \|Y\| = 0 \end{cases}$$

### Demostración por inducción

- $\mu(\emptyset) = 0$ , se cumple por definición de la función  $\mu$  cuando  $\|Y\| = 0$ ,
- Si  $A_1, A_2, \dots$  son disyuntos entonces  $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ ,  $\forall A_i \in \gamma$
- Mostrar para  $\|Y\| = 1$ ,  $\mu(A_1) = \frac{g(y_1)}{h(y_1)}$ ,  $A_1 \in Y$ .  $\mu(\bigcup_{n=1}^1 A_n) = \mu(A_1) = \frac{g(y_1)}{h(y_1)} = \sum_{n=1}^1 \mu(A_n)$ ,  $\forall A_i \in \gamma$
- Suponer para  $\|Y\| = m$ ,  $\mu(\bigcup_{n=1}^m A_n) = \sum_{n=1}^m \mu(A_n)$ ,  $\forall A_i \in \gamma$  y demostrar para  $\|Y\| = m+1$ . Por hipótesis los elementos  $A_1, A_2, \dots$  son disyuntos entonces:

$$\begin{aligned} \mu(\bigcup_{n=1}^{m+1} A_n) &= \mu(\bigcup_{n=1}^m A_n \cup A_{m+1}) = \mu(\bigcup_{n=1}^m A_n) + \mu(A_{m+1}) = \sum_{n=1}^m \mu(A_n) + \mu(A_{m+1}) = \sum_{n=1}^m \frac{g(y_n)}{h(y_n)} + \frac{g(y_{m+1})}{h(y_{m+1})} \\ &= \sum_{n=1}^{m+1} \frac{g(y_n)}{h(y_n)} = \sum_{n=1}^{m+1} \mu(A_n) \end{aligned}$$

Se define la función de disimilaridad  $d$  sobre un par de elementos de  $\gamma$  como una aplicación de  $d: X \times X \rightarrow [0, +\infty)$  así:  $d(x, y) = |\mu(x) - \mu(y)|$ ,  $\forall x, y \in \gamma$ .

Pero esta función debe satisfacer las siguientes propiedades,  $\forall x, y, z \in \gamma$ .

- $d(x, y)$ , es un número positivo o cero.  
 $d(x, y) = |\mu(x) - \mu(y)|$ , y por definición del valor absoluto la propiedad se satisface.
- $d(x, y) = d(y, x)$   
 $d(x, y) = |\mu(x) - \mu(y)|$ , por conmutatividad de la resta en el valor absoluto  $|\mu(y) - \mu(x)| = d(y, x)$ .
- $d(x, x) = 0$ ,  $d(x, x) = |\mu(x) - \mu(x)| = 0$ .
- $d(x, y) \leq d(x, z) + d(z, y)$ ,  $d(x, y) = |\mu(x) - \mu(y)| = |\mu(x) - \mu(y) + \mu(z) - \mu(z)|$   
 $= |(\mu(x) - \mu(z)) - (\mu(z) - \mu(y))| \leq |\mu(x) - \mu(z)| + |\mu(z) - \mu(y)| = d(x, z) + d(z, y)$   
 $\therefore d(x, y) \leq d(x, z) + d(z, y)$ ,

En la demostración se aplicó una de las propiedades del valor absoluto:  $|x + y| \leq |x| + |y|$ , también llamada desigualdad triangular, la cual se demuestra usando otras propiedades del valor absoluto:

- $-|x| \leq x \leq |x|$ ,  $\forall x \in \mathfrak{R}$ , (propiedad del valor absoluto)
- $-|y| \leq y \leq |y|$ ,  $\forall y \in \mathfrak{R}$ ,

c.  $(-|x|) + (-|y|) \leq x + y \leq |x| + |y|$ , (sumando a y b),

d.  $-(|x| + |y|) \leq x + y \leq (|x| + |y|)$ , aplicando la propiedad  $-|a| \leq x \leq |a| \Leftrightarrow |x| \leq a \quad \therefore |x + y| \leq |x| + |y|$

## B. Método de clasificación

### 1) Condiciones iniciales

- $R = \emptyset$ , el repositorio debe estar totalmente vacío.
- $T \neq \emptyset$ , deben estar definidas algunas categorías de entidad.
- $m \neq \emptyset, \forall m \in t, \forall t \in T$ , toda categoría de entidad debe tener un nombre definido.
- $n = 0, \forall n \in t, \forall t \in T$ , toda categoría de entidad por la primera condición inicial, no debe tener objetos de aprendizaje asociados.
- $V \neq \emptyset, \forall V \in t, \forall t \in T$ , toda categoría de entidad debe tener definidos identificadores.
- $\forall a, b \in V, a \neq b, \forall V \in t, \forall t \in T$ , todos los indicadores deben ser diferentes, se define la comparación de elementos de  $V$  como la comparación de sus identificadores ( $i$ ).
- $v = 0, \forall n \in V, \forall V \in t, \forall t \in T$ , el valor de todos los indicadores de las categorías de entidad debe ser cero.

### 2) Condiciones para adicionar un objeto $O$

- Identifique a que categoría de entidad  $t$  corresponde el objeto de aprendizaje  $O$  a adicionar.
- Obtenga todos los resultados de la medición de los indicadores definidos en  $t$  para  $O$ , considerando  $C$  y  $S$ , en donde  $C, S \in O$  y construya apropiadamente  $P$  con los resultados, ahora complete la terna  $\langle C, S, P \rangle$ .
- Adicione  $O$  al repositorio  $R$  y actualice apropiadamente el componente  $V$  de  $t$  (sume los valores obtenidos en  $P$  en el punto anterior a los indicadores correspondientes).

### 3) Condiciones para retirar un objeto $O$

- Identifique a que categoría de entidad  $t$  corresponde el objeto de aprendizaje  $O$  a adicionar.
- Retire  $O$  del repositorio  $R$  y actualice apropiadamente el componente  $V$  de  $t$  (restando los valores en  $P$  en el punto anterior a los indicadores correspondientes).

### 4) Clasificación de los objetos de aprendizaje $O$

Construya una secuencia ordenada  $L$ , tal que  $\forall x \in R, x \in L$ , y que  $\forall l_i, l_j \in L$  en donde  $i < j$  se debe cumplir que  $\mu(l_i) \leq \mu(l_j)$ .

Una vez construida la secuencia ordenada  $L$ , los peores objetos de aprendizaje pertenecientes al repositorio  $R$  serán los ubicados en las primeras posiciones de la secuencia y por tanto serán los mejores candidatos a entrar a un proceso de mejoramiento. Una vez se ha decidido mejorar un objeto de aprendizaje, éste deberá ser retirado del repositorio, mejorado y vuelto a adicionar al mismo una vez se hayan hecho las mejoras.



Como se puede observar, este algoritmo no considera directamente el contenido del ( $C$ ) objeto de aprendizaje o su metadato ( $S$ ), sino el conjunto de todos los pares ordenados  $(i, v_i)$ , llamado  $P$ , en donde  $i$  es el código que identifica al indicador y  $v_i$  es el valor asociado a él. Por supuesto para poder construir a  $P$  es necesario que un personal capacitado evalúe a  $C$  y/o a  $S$ . En este orden de ideas, este algoritmo puede ser clasificado como un algoritmo que utiliza un método basado en restricciones, pues todos los objetos de aprendizaje que tengan el mismo número de indicadores por encima del promedio (inclusive) y sin importar cuales, pertenecerán a la misma categoría.

Como ya es sabido, el mejor algoritmo de ordenamiento basado en comparaciones tiene una complejidad computacional del orden de  $O(n \ln n)$ , como por ejemplo el QuickSort, sin embargo, para este caso particular, se puede utilizar un algoritmo de ordenamiento de orden  $O(n)$ , llamado RadixSort, usándolo para ordenar  $V[\mu(\circ) * 10]$ , en donde 10 restringe a una precisión de una cifra decimal (una precisión más que suficiente para esta aplicación). La explicación de este algoritmo y su implementación se sale del alcance de este artículo, pero puede ser consultado en la literatura de ciencias de la computación, como por ejemplo el libro de Thomas H. Cormen [8] "Introduction to Algorithms", especialmente en su capítulo "Sorting in Linear Time".

Las aglomeraciones se pueden obtener al identificar los objetos de aprendizaje (entidades) que tiene igual porcentaje de aceptación, o en un defecto tengan un grado de disimilaridad pequeño (parámetro del usuario).

##### 5) *Requerimientos sobre el algoritmo*

Un algoritmo de aglomeramiento debe cumplir una serie de requerimientos:

- **Escalabilidad:** El algoritmo planteado no tiene problema en trabajar con un número alto de objetos, más aún si utiliza un algoritmo de ordenamiento en tiempo lineal.
- **Trabajar con diferentes tipos de datos:** Esta primera investigación restringe a indicadores que puedan ser valorados numéricamente. Investigaciones futuras deberán ampliar el abanico en los tipos de datos utilizados para valorar los indicadores.
- **Descubrir agrupaciones en formas arbitrarias:** Este algoritmo utiliza la Métrica Manhattan (o del taxista), la cual se reduce en este caso a  $d(x, y) = |\mu(x) - \mu(y)|$ . Este algoritmo toma los objetos de aprendizaje y los ubica en una recta que representa el intervalo  $[0.0; 1.0]$ , en esta recta se ubican puntos que representan porcentajes de aceptación de un objeto de aprendizaje  $\frac{g(y_n)}{h(y_n)}$ , los cuales coinciden con las aglomeraciones realizadas. Como podrá visualizarse en la Sección IV-B, los objetos pueden moverse de una aglomeración a otra, cada vez que se realice una operación sobre el repositorio  $R$ , por tanto y bajo la óptica de los objetos de aprendizaje, sería más apropiado hablar de que el algoritmo descubre formas arbitrarias.

- **Mínima cantidad de parámetros de entrada:** Como ya se ha dicho, este algoritmo requiere por cada objeto de aprendizaje  $O$ , un conjunto  $P$  de todos los pares ordenados  $(i, v_i)$  en donde  $i$  es el código que identifica al indicador y  $v_i$  es el valor asociado a él.
- **Habilidad para trabajar datos con ruido:** Si un objeto está mal evaluado, esta evaluación errónea puede (y seguramente lo hará) modificar la clasificación de los demás objetos de esa categoría, pues puede alterar los valores almacenados en la componente  $v$  de la categoría  $t$  a la que pertenece el objeto de aprendizaje mal evaluado. La ventaja es que una vez detectada la inconsistencia, este objeto puede ser retirado del repositorio, restaurando los valores anteriores de  $v$  y haciendo que se restaure la clasificación original.
- **Aglomeración incremental e insensibilidad a la cantidad de objetos de entrada:** Como se pudo observar en la Sección III-B2 y III-B3, es muy sencillo insertar o retirar objetos de aprendizaje del repositorio  $R$ . Esto por supuesto no trae nuevas complicaciones a la hora de crear la secuencia ordenada  $L$ , pues lo que hay que hacer es nuevamente ordenar los objetos de repositorio usando preferiblemente un algoritmo de tiempo lineal como el ya expuesto.
- **Alta dimensionalidad:** Este algoritmo puede trabajar con objetos de aprendizaje con un número elevado de indicadores (norma de  $P$ ), es más, puede trabajar con diferentes categorías de objetos de aprendizaje, cada uno con una cantidad diferente de indicadores. La dificultad entonces no estará en el algoritmo, sino en el equipo que debe evaluar cada objeto de aprendizaje para construir a  $P$ .
- **Agrupación basada en restricciones:** Este algoritmo crea aglomeraciones que satisfagan ciertas restricciones, en este caso el porcentaje de aceptación del objeto de aprendizaje según la valoración de un conjunto de indicadores definidos para cada categoría de objetos.
- **Interpretabilidad y usabilidad:** No debe haber discusión en que los resultados de las aglomeraciones son fácilmente interpretados, comprendidos y que pueden ser usables. En este caso, los diferentes objetos (sin importar a la clase a la que pertenezcan), van a estar clasificados en aglomeraciones que representan el porcentaje de aceptación de esos objetos de aprendizaje, tal y como ya se ha explicado en reiteradas ocasiones. La usabilidad de la información radica principalmente en la detección de los mejores candidatos a entrar a un proceso de mejoramiento.

#### IV. EJEMPLARIZANDO – CONTEXTO CONCRETO

En esta sección se intentará definir un contexto para poder crear un ejemplo en la Sección IV-B.

Existe un riesgo al definir un contexto concreto y es que el lector desvie su atención a problemas que se encuentran fuera del alcance de este artículo (ej: la formulación apropiada de indicadores, la identificación de conceptos medibles o de atributos significativos, entre muchos otros problemas). En este orden de ideas, tener un contexto concreto permite que el lector se focalice en comprender el mecanismo de medición definido en esta investigación.

Los atributos utilizados son tomados de la página web “TECNOLOGIA EDUCATIVA - WEB PERE MARQUÈS”[23], al igual que las categorías de entidad (Videos didácticos<sup>5</sup>, Transparencias y Diapositivas<sup>6</sup>, Aplicaciones multimedia<sup>7</sup>, Páginas Web<sup>8</sup>).

- **Modelo de calidad:**

<i>“Modelo de calidad para objetos de aprendizaje”</i>
--

- **Conceptos medibles** para el “Modelo de calidad para objetos de aprendizaje”, según las necesidades de información que se tengan:

<i>“Aspectos funcionales”</i>	<i>“Aspectos Técnicos”</i>
<i>“Aspectos Pedagógicos”</i>	

- **Categorías de entidad** para el “Modelo de calidad para objetos de aprendizaje”

<i>“Video didácticos”</i>
<i>“Aplicaciones multimedia educativas”</i>
<i>“Transparencias y Diapositivas”</i>
<i>“Páginas Web Educativas”</i>

- **Atributos por cada categoría de entidad** del “Modelo de calidad para objetos de aprendizaje”

- “Video didácticos”

$A_{11}$	<i>“Eficacia (puede facilitar el logro de sus objetivos)”</i>
$A_{12}$	<i>“Relevancia curricular de los objetivos que persigue”</i>
$A_{13}$	<i>“Contenidos (calidad, profundidad, organización)”</i>
$A_{14}$	<i>“Estructura y ritmo (guión claro, secuenciación...)”</i>
$A_{15}$	<i>“Planteamiento audiovisual (interacción)”</i>
$A_{16}$	<i>“Textos, gráficos y animaciones”</i>
$A_{17}$	<i>“Capacidad de motivación (atractivo, interés)”</i>
$A_{18}$	<i>“Adecuación al usuario (contenidos, actividades)”</i>
$A_{19}$	<i>“Planteamiento didáctico (organizadores, resumen...)”</i>

- “Transparencias y Diapositivas”

$A_{21}$	<i>“Eficacia. El material facilita la comprensión”</i>
$A_{22}$	<i>“Contenidos. Los contenidos son significativos, correctos, actuales y de calidad”</i>
$A_{23}$	<i>“Adecuación de los contenidos a los destinatarios”</i>

- “Aplicaciones multimedia educativas”

$A_{31}$	<i>“Versatilidad didáctica: modificable, niveles, informes.”</i>
$A_{32}$	<i>“Documentación, guía didáctica”</i>
$A_{33}$	<i>“Interacción (tipo de diálogo, análisis respuestas)”</i>
$A_{34}$	<i>“Estructura y navegación por las actividades, metáforas”</i>
$A_{35}$	<i>“Potencialidad de los recursos didácticos (actividades,)”</i>
$A_{36}$	<i>“Posibilita el trabajo cooperativo, da facilidades”</i>

<sup>5</sup> <http://www.pangea.org/peremarques/videoav2.htm>

<sup>6</sup> <http://www.pangea.org/peremarques/diapoeva.htm>

<sup>7</sup> <http://www.pangea.org/peremarques/evalua.htm>

<sup>8</sup> <http://www.pangea.org/peremarques/evalweb.htm>

- “Páginas Web Educativas”

A <sub>41</sub>	“Carácter multilingüe, al menos algunos apartados”
A <sub>42</sub>	“Relevancia, interés de los contenidos y servicios”
A <sub>43</sub>	“Ejecución fiable, velocidad de acceso adecuada”
A <sub>44</sub>	“Fomento del autoaprendizaje. toma decisiones”

- **Atributos relacionados con cada concepto medible** del “Modelo de calidad para objetos de aprendizaje”

- “Aspectos Funcionales”

A <sub>11</sub>	A <sub>12</sub>	A <sub>21</sub>	A <sub>31</sub>	A <sub>32</sub>	A <sub>41</sub>	A <sub>42</sub>
-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

- “Aspectos Técnicos”

A <sub>13</sub>	A <sub>14</sub>	A <sub>15</sub>	A <sub>16</sub>	A <sub>22</sub>	A <sub>33</sub>	A <sub>34</sub>	A <sub>43</sub>
-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

- “Aspectos Pedagógicos”

A <sub>17</sub>	A <sub>18</sub>	A <sub>19</sub>	A <sub>23</sub>	A <sub>35</sub>	A <sub>36</sub>	A <sub>44</sub>
-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

- **Entidades (objetos de aprendizaje) por cada categoría de entidad**

- “Video didácticos”

“Video 1”	“Video 2”	“Video 3”
“Video 4”	“Video 5”	“Video 6”
“Video 7”	“Video 8”	

- “Transparencias y Diapositivas”

“Presentación 1”	“Presentación 2”	“Presentación 3”
“Presentación 4”	“Presentación 5”	“Presentación 6”
“Presentación 7”		

- “Aplicaciones multimedia educativas”

“Multimedia 1”	“Multimedia 2”	“Multimedia 3”
“Multimedia 4”	“Multimedia 5”	“Multimedia 6”
“Multimedia 7”	“Multimedia 8”	“Multimedia 9”

- “Páginas Web Educativas”

“Web 1”	“Web 2”	“Web 3”
“Web 4”	“Web 5”	“Web 6”

Una vez definido por cada atributo su forma de medición (Método de medición, Función de cálculo, o por un modelo de análisis) se procede a su medición por parte de un equipo competente, obteniendo así unos resultados de medición por cada entidad (objeto de aprendizaje), según la categoría a la que pertenezca.

- “Video didácticos”

	A <sub>11</sub>	A <sub>12</sub>	A <sub>13</sub>	A <sub>14</sub>	A <sub>15</sub>	A <sub>16</sub>	A <sub>17</sub>	A <sub>18</sub>	A <sub>19</sub>
“Video 1”	73	47	78	72	91	97	76	77	54
“Video 2”	89	72	44	17	27	91	76	82	93
“Video 3”	45	19	41	89	79	97	92	17	9
“Video 4”	11	71	34	97	3	98	78	10	63
“Video 5”	29	66	71	52	13	96	61	67	32
“Video 6”	72	2	32	48	54	44	40	8	100
“Video 7”	12	2	18	53	4	90	52	33	50
“Video 8”	48	4	70	22	32	48	53	61	21

Se sobreentiende que cada atributo tiene definida una medida (medida base, medida derivada o indicador), la cual debe tener definida una unidad de medición, una escala de medición (ej: 0 a 100) y esta última debe pertenecer a un tipo específico de escala.

- “Transparencias y Diapositivas”

	$A_{21}$	$A_{22}$	$A_{23}$
“Presentación 1”	61	34	36
“Presentación 2”	88	94	71
“Presentación 3”	76	45	1
“Presentación 4”	40	14	29
“Presentación 5”	92	20	35
“Presentación 6”	23	21	85
“Presentación 7”	50	78	38

- “Aplicaciones multimedia educativas”

	$A_{31}$	$A_{32}$	$A_{33}$	$A_{34}$	$A_{35}$	$A_{36}$
“Multimedia 1”	90	21	82	67	55	66
“Multimedia 2”	6	0	61	25	46	56
“Multimedia 3”	46	34	0	36	40	53
“Multimedia 4”	83	82	67	24	79	88
“Multimedia 5”	71	5	49	45	94	41
“Multimedia 6”	91	97	2	7	98	57
“Multimedia 7”	57	53	16	90	51	53
“Multimedia 8”	73	86	46	92	43	76
“Multimedia 9”	69	92	20	5	49	8

- “Páginas Web Educativas”

	$A_{31}$	$A_{32}$	$A_{33}$	$A_{34}$
“Web 1”	7	39	82	80
“Web 2”	38	62	76	84
“Web 3”	3	49	16	68
“Web 4”	49	17	91	70
“Web 5”	53	79	58	1
“Web 6”	24	11	81	18

En este punto es bueno mencionar que a cada categoría de entidad por separado se le puede aplicar un método de clasificación, sin embargo, cuando las categorías entidad tienen un número diferente de atributos, no es posible aplicar estos métodos a todas las categorías simultáneamente.

#### A. Mecanismo de valoración de indicadores

Lo ideal es que la institución disponga de un servicio web que le permita, preferiblemente, a un equipo de personas competentes medir cada uno de los objetos que van a ingresar al repositorio. Esto quiere decir que los autores no ingresan directamente al repositorio sus objetos, sino que los envían primero al dicho equipo. Otra estrategia, pero menos confiable, es que cada autor de un objeto de aprendizaje debe autoevaluar su aporte antes de adicionarlo al repositorio.

Por otro lado, la institución también debe disponer de un equipo de mejoramiento que ejecute el algoritmo de clasificación de objetos cada vez que lo requiera para identificar los mejores candidatos a ser mejorados (los de menor porcentaje de aceptación), una vez hecha esta labor, los retira del repositorio, los mejora y los ingresa nuevamente. Estas actividades de retirar e ingresar objetos al repositorio afectan, como ya se ha dicho, muy probablemente las aglomeraciones encontradas antes de realizar estas actividades.

## B. Ejemplo del método de clasificación

### 1) Condiciones iniciales

- $R = \emptyset$ , el repositorio debe estar totalmente vacío.
- Categorías de entidad

```
T = { ( "Video didácticos", 0,
      { (11, A11, 0), (12, A12, 0), (13, A13, 0), (14, A14, 0), (15, A15, 0), (16, A16, 0),
        (17, A17, 0), (18, A18, 0), (19, A19, 0) }
    ),
  ( "Transparencias y Diapositivas", 0, { (21, A21, 0), (22, A22, 0), (23, A23, 0) } ),
  ( "Aplicaciones multimedia educativas", 0,
    { (31, A31, 0), (32, A32, 0), (33, A33, 0), (34, A34, 0), (35, A35, 0), (36, A35, 0) }
  );
  ( "Páginas Web Educativas", 0, { (41, A41, 0), (42, A42, 0), (43, A43, 0), (44, A44, 0) } )
}
```

### 2) Objetos de aprendizaje

Un Objeto de aprendizaje, llamado *Entidad* en el contexto, se define como la terna  $\langle C, S, P \rangle$  en donde  $C$  es el contenido del objeto de aprendizaje,  $S$  es el conjunto de los valores que toman los atributos definidos en el metadato asociado al objeto y  $P$  es el conjunto de todos los pares ordenados  $(i, v_i)$  en donde  $i$  es el código que identifica al indicador y  $v_i$  es el valor asociado a él. Para no repetir innecesariamente los resultados de la evaluación de todos los objetos de aprendizaje, se va a ilustrar sólo para el objeto de aprendizaje llamado "Presentación 1".

$$\text{"Presentación 1"} = \langle C, S, \{(21;61); (22;34); (23;36)\} \rangle$$

### 3) Repositorio de Objetos de aprendizaje

Una vez creados los objetos de aprendizaje (Entidades), estos se insertan en el repositorio  $R$ .

```
R = { "Video 1", "Video 2", "Video 3", "Video 4", "Video 5", "Video 6", "Video 7", "Video 8",
      "Presentación 1", "Presentación 2", "Presentación 3", "Presentación 4", "Presentación 5",
      "Presentación 6", "Presentación 7", "Multimedia 1", "Multimedia 2", "Multimedia 3",
      "Multimedia 4", "Multimedia 5", "Multimedia 6", "Multimedia 7", "Multimedia 8",
      "Multimedia 9", "Web 1", "Web 2", "Web 3", "Web 4", "Web 5", "Web 6"
    }
```

Una vez insertados todos los objetos de aprendizaje tal y como se indica en la Sección III-B2, las categorías de entidad quedan así:

```
T = { ( "Video didácticos", 8,
      { (11, A11, 379), (12, A12, 283), (13, A13, 388), (14, A14, 450), (15, A15, 303),
        (16, A16, 661), (17, A17, 528), (18, A18, 355), (19, A19, 422) }
    );
  ("Transparencias y Diapositivas", 7, {(21, A21, 430), (22, A22, 306), (23;A23; 295)});
  ( "Aplicaciones multimedia edu", 9,
    { (31,A31,586), (32,A32,510), (33,A33,343), (34,A34,391), (35,A35,555), (36,A35,498) } ),
  ( "Páginas Web Educativas", 6, {(41,A41,174), (42,A42,257), (43,A43,404), (44,A44,321) } )
}
```

#### 4) Aglomeración de Objetos de aprendizaje

Una vez se tiene lleno el repositorio, se procede construir la secuencia ordenada  $L$ , tal que  $\forall x \in R, x \in L$  y que  $\forall l_i, l_j \in L$  en donde  $i < j$  se debe cumplir que  $\mu(l_i) \leq \mu(l_j)$ . Para este caso particular se trabajó con una cifra decimal, lo que produce máximo 11 aglomeraciones.

$$\{ \mu(\cdot) = 0.0, \mu(\cdot) = 0.1, \mu(\cdot) = 0.2, \mu(\cdot) = 0.3, \mu(\cdot) = 0.4, \mu(\cdot) = 0.5, \\ \mu(\cdot) = 0.6, \mu(\cdot) = 0.7, \mu(\cdot) = 0.8, \mu(\cdot) = 0.9, \mu(\cdot) = 1.0 \}$$

Donde  $(\cdot)$  representa cualquier miembro de la aglomeración. Entre más cifras decimales considere, mayor será el número de aglomeraciones y más fina será la clasificación de los objetos. Para este caso teórico y para muchos casos prácticos, una cifra decimal es más que suficiente.

$\mu(\cdot) = 0.0$	"P/ción 1"	$\mu(\cdot) = 0.5$	"Web 1"
	"P/ción 4"		"Web 3"
	"M/media 3"		"Web 5"
$\mu(\cdot) = 0.1$	"Video 7"	$\mu(\cdot) = 0.6$	"Video 4"
	"M/media 7"		"M/medi 4"
			"M/medi 8"
$\mu(\cdot) = 0.2$	"M/media 2"	$\mu(\cdot) = 0.7$	"Video 2"
	"M/media 9"		"P/ción 3"
$\mu(\cdot) = 0.3$	"Web 6"	$\mu(\cdot) = 0.8$	"Web 4"
	"Video 6"		
	"Video 8"	$\mu(\cdot) = 1.0$	"Video 1"
	"P/ción 5"		"P/ción 2"
	"P/ción 6"		"Web 2"
"P/ción 7"			
$\mu(\cdot) = 0.4$	"Video 3"		
	"Video 5"		
	"M/media 1"		
	"M/media 5"		
	"M/media 6"		

La forma de interpretar esta clasificación, se encuentra en el valor de  $\mu(\cdot)$ , pues este valor indica el porcentaje de aceptación. En tal caso,  $\{ \text{"Presentación 1"}, \text{"Presentación 4"}, \text{"Multimedia 3"} \}$  son los peores objetos de aprendizaje del repositorio, pues tiene un porcentaje de aceptación de cero. Esto quiere decir que ninguno de sus indicadores se encuentra por encima del promedio (inclusive) de este indicador en la categoría de entidad a la que corresponde el objeto de aprendizaje. Observe que en esta aglomeración hay objetos de diferentes clases.

Como se puede deducir, los objetos de aprendizaje  $\{ \text{"Video 1"}, \text{"Multimedia 2"}, \text{"Web 2"} \}$ , son los mejores (hasta ahora) en el repositorio, no quiere decir que no tengan problemas, lo que quiere decir es que sus indicadores sí se encuentran por encima del promedio (o son iguales) de este indicador en la categoría de entidad a la que corresponde el objeto de aprendizaje.

Ahora suponga que el objeto de aprendizaje "Multimedia 3" es retirado del repositorio, entonces la pregunta es: ¿Cómo afecta esto las aglomeraciones?, y la respuesta es, de ninguna forma (se deja al lector como ejercicio). Pero una vez el objeto de aprendizaje "Multimedia 3\*" es mejorado, él es medido e ingresado al repositorio como:

“Multimedia 3\*” =  $\langle C_{33}, S_{33}, \{ (31; 97); (32; 55); (33; 81); (34; 63); (35; 76); (36; 69) \} \rangle$

$\mu(\cdot) = 0.0$	“P/ción 1”	$\mu(\cdot) = 0.5$	“Web 1”
	“P/ción 4”		“Web 3”
$\mu(\cdot) = 0.1$	“Video 7”	$\mu(\cdot) = 0.6$	“Web 5”
	“M/media 2”		“Video 4”
	“M/media 7”		“M/med 3*”
	“M/media 9”		“M/medi 4”
$\mu(\cdot) = 0.3$	“Web 6”	$\mu(\cdot) = 0.7$	“Video 2”
	“Video 6”		“P/ción 3”
	“Video 8”	$\mu(\cdot) = 0.8$	“Web 4”
	“P/ción 5”		
	“P/ción 6”		
	“P/ción 7”		
	“M/media 5”		
“M/media 6”			
$\mu(\cdot) = 0.4$	“Video 3”	$\mu(\cdot) = 1.0$	“Video 1”
	“Video 5”		“P/ción 2”
	“M/media 1”		“Web 2”
	“M/media 8”		

Como se puede observar, este pequeño cambio en la inserción de un nuevo objeto de aprendizaje “Multimedia 3\*”, provocó cambio en las aglomeraciones:

- Desaparece la aglomeración asociada a  $\mu(\cdot) = 0.2$ .
- Los objetos de aprendizaje “Multimedia 2”, “Multimedia 9”, “Multimedia 5”, “Multimedia 6” “Multimedia 8” y “Multimedia 3\*” cambian de aglomeración.

Entidad	$\mu(\cdot)$ anterior	$\mu(\cdot)$ actual
“Multimedia 2”	0.2	0.1
“Multimedia 9”	0.2	0.1
“Multimedia 5”	0.4	0.3
“Multimedia 6”	0.4	0.3
“Multimedia 8”	0.6	0.4
“Multimedia 3*”	0.0	0.6

Como se pudo ilustrar con el ejemplo, cambios en el repositorio pueden afectar las aglomeraciones de los objetos de aprendizaje. Sin embargo, también se puede observar que estos cambios, si los hay, se producen únicamente en la categoría o categorías que fueron afectadas con los cambios en el repositorio. Estos cambios pueden ser significativos como por ejemplo para el objeto de aprendizaje llamado “Multimedia 3”



## V. CONCLUSIÓN

- Se logró desarrollar un mecanismo automático de clasificación de un conjunto de objetos de aprendizaje según el valor de unos indicadores.
- Se construyó un espacio de medida (la terna (Conjunto de objetos  $X$ , la sigma álgebra  $\Sigma$ , una función de medida  $\mu$ )) que sirve para clasificar automáticamente un conjunto de objetos de aprendizaje según el valor numérico de unos indicadores (Sección III-A).
  - Se logró demostrar matemáticamente que la función de medida  $\mu$  definida es función válida en la Sección III-A.
  - Se logró demostrar matemáticamente que la función de disimilaridad definida satisface todas las propiedades en la Sección III-A.
  - Se definió un método de clasificación que utilice la medida definida (Sección III-B).
  - Se analizaron los requerimientos que deben cumplir los algoritmos de aglomeramiento con respecto a la nueva propuesta y se pudieron observar sus bondades.
  - Se determinaron los indicadores para clasificar los objetos de aprendizaje en un contexto particular (Sección IV).
  - Se estableció un mecanismo de valoración de los indicadores definidos (Sección IV-A).
  - Se desarrolló un ejemplo del método de clasificación para el contexto particular definido (Sección IV-B).
- Para este mecanismo de clasificación en particular, se consideran sólo los indicadores que puedan ser valorados numéricamente. Sin embargo, esto no debe ser visto como una debilidad, por el contrario, debe ser visto como una oportunidad para futuras investigaciones. La idea, sería crear una nueva función de medida, que por supuesto hay que demostrar matemáticamente, tal y como se hizo con la función propuesta.
- Como se pudo mostrar, es posible construir mecanismos automáticos de clasificación de un conjunto de objetos de aprendizaje. Aún si estos objetos son de categorías diferentes.

## REFERENCIAS

- [1] Anderberg, Michael R. (1973), Cluster analysis for applications: Academic Press.
- [2] Berkhin Pavel (2002). Survey of clustering data mining techniques. Consultado en marzo 2009, en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.3739>.
- [3] Boneu Josep M., Plataformas abiertas de e-learning para el soporte de contenidos educativos abiertos. Consultado en abril de 2009 en <http://www.uoc.edu/rusc/4/1/dt/esp/boneu.pdf>.
- [4] C. M Benjamin, Ke Fung y Ester Wang Martin. (2003), Large hierarchical document clustering using frequent itemsets, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.9326>.
- [5] Cao Zhe, Qin Tao, Liu Tie-Yan, Tsai Ming-Feng y Li Hang. (2007), "Learning to Rank: From Pairwise Approach to Listwise Approach", Consultado en octubre de 2009 en <http://research.microsoft.com/en-us/people/tyliu/139.pdf>.
- [6] Chakrabarti Soumen. (2000), Data mining for hypertext: A tutorial survey, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.4297>.
- [7] Chechik Gal, Sharma Varun, Shalit Uri y Bengio Samy (2009), "Large Scale Online Learning of Image Similarity Through Ranking", Consultado en octubre de 2009 en <http://snowbird.djvuzone.org/abstracts/119.pdf>.
- [8] Cormen Thomas H., Leiserson Charles E., Rivest Ronald L. y Stein Clifford (2001), Introduction to Algorithms, Second Edition: The MIT Press, McGraw-Hill.
- [9] Cui Xiaohui y Potok Thomas E. (2005), Document Clustering Analysis Based on Hybrid PSO+K-means Algorithm, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.127.8448>.
- [10] Driver Bruce K. (2003), Analysis Tools with Applications (25-26 y 131-135): Springer.
- [11] Deshpande Mukund y Karypis George (2002), Using Conjunction of Attribute Values for Classification, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.16.6905>.
- [12] desJardins Marie, Eaton Eric y Wagstaff Kiri L. (2006), "Learning User Preferences for Sets of Objects", Consultado en octubre de 2009 en <http://maple.cs.umbc.edu/~ericeaton/papers/desJardins2006Learning.pdf>.
- [13] Ferreira Mateus, García Félix, Ruiz Francisco, Bertoa Manuel F., Calero Coral, Vallecillo Antonio, Piattini Mario y Mora Beatriz. (2006), Medición del Software Ontología y Metamodelo (Informe Técnico UCLM-TSI-001), Departamento de Tecnologías y Sistemas de Información, Universidad de Castilla-La Mancha y Universidad de Málaga, España, , Consultado en abril de 2009 en <http://www.esi.uclm.es:8080/tsi/informes/UCLM-TSI-001.pdf>.
- [14] Graf Sabine y List Beate (2005), "An Evaluation of Open Source E-Learning Platforms Stressing Adaptation Issues", Consultado en octubre de 2009 en <http://wit.tuwien.ac.at/people/list/publications/icalt2005.pdf>.
- [15] Han Jiawei y Kamber Micheline (2006), Data Mining, Concepts and Techniques, Second Edition: Morgan Kaufmann Publishres, Chapter 7 "Cluster Analysis", Pag. 383-464
- [16] Hammouda Khaled M. , Member Student y Kamel Mohamed S. (2004), Efficient phrase-based document indexing for Web document clustering, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.6184>.
- [17] Hofmann Thomas y Puzicha Jan. (2000), Data mining for hypertext: A tutorial survey, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.97>.
- [18] IEEE LOM Institute for Electrical and Electronic Engineers Standards Department, Learning Object Metadata (IEEE 1481.12.1-2002), Consultado en mayo 2008 en <http://ieeexplore.ieee.org/xpl/standards.jsp?findtitle=learning+object+metadata&letter=learning+object+metadata&imageField.x=0&imageField.y=0&opentree=on>, o también se puede acceder en español en [http://www-gist.det.uvigo.es/~lanido/LOMes/LOMv1\\_0\\_Spanish.pdf](http://www-gist.det.uvigo.es/~lanido/LOMes/LOMv1_0_Spanish.pdf)

- [19] Kaur Kamaldeep, Kaur Arvinder y Malhotra Ruchika. (2009), "Alternative Methods to Rank the Impact of Object Oriented Metrics in Fault Prediction Modeling using Neural Networks", Consultado en octubre 2009 en <http://www.waset.org/journals/waset/v19/v19-27.pdf>.
- [20] Khan Inayatullah, Saffari Amir y Bischof Horst. (2009), "TVGraz: Multi-Modal Learning of Object Categories by Combining Textual and Visual Features", Consultado en octubre 2009 en <http://oagm2009.icg.tugraz.at/papers/p36.pdf>.
- [21] Li Hang. (2008), "Learning to Rank: Problem, Challenge, and Opportunity", Microsoft Research Asia, Consultado en octubre 2009 en <http://research.microsoft.com/users/tyliu/>.
- [22] Liu Tie-Yan. (2008), "Learning to Rank: from Pairwise Approach to Listwise Approach", Microsoft Research Asia, Consultado en octubre 2009 en <http://research.microsoft.com>.
- [23] Marquès Pere. (2009), Sitio Web sobre Tecnología Educativa, Consultado en julio 2009 en <http://www.pangea.org/peremarques/>.
- [24] Mimaroglu Selim, Simovici Dan A., "Bit Sequences and Biclustering of Text Documents", , Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.4386>.
- [25] Nesbit John C. y Li Jerry. (2004), "Web-Based Tools for Learning Object Evaluation" , Consultado en octubre de 2009 en [http://www.sfu.ca/~jzli/publications/Nesbit\\_Li\\_2004.pdf](http://www.sfu.ca/~jzli/publications/Nesbit_Li_2004.pdf).
- [26] Nesbit John C., Li Jerry, Leacock Tracey L. (2004), "Web-Based Tools for Collaborative Evaluation of Learning Resources", Consultado en octubre 2009 en [http://www.iiisci.org/journal/CV\\$/sci/pdfs/P843017.pdf](http://www.iiisci.org/journal/CV$/sci/pdfs/P843017.pdf).
- [27] Nie Zaiqing, Zhang Yuanzhi, Wen Ji Rong y Ma WeiYing. (2005), "Object Level Ranking: Bringing Order to Web Objects", Consultado en octubre 2009 en <http://www2005.org/cdrom/docs/p567.pdf>.
- [28] "Proceedings of the NIPS 2005 Workshop on Learning to Rank", Whistler, BC, Canada, (2005), Consultado en octubre 2009 en <http://web.mit.edu/shivani/www/Ranking-NIPS-05/Proceedings/proceedings-nips05workshop-ranking.pdf>.
- [29] Ochoa Xavier (2006), The research about Metrics for Metadata, Hypermedia and Databases Group (HMDB) of the Katholieke Universiteit Leuven (KULeuven), Consultado en marzo de 2009 en <http://ariadne.cti.espol.edu.ec/M4M/index.html>.
- [30] Ochoa Xavier y Duval Erik (2006), "Metrics for Learning Object Metadata" , Consultado en octubre de 2009 en <http://ariadne.cti.espol.edu.ec/xavier/papers/Ochoa-DC-ECTEL2006.pdf>.
- [31] Ochoa Xavier y Duval Erik (2006), "Use of contextualized attention metadata for ranking and recommending learning objects", Consultado en octubre de 2009 en [http://portal.acm.org/ft\\_gateway.cfm?id=1183608&type=pdf&coll=GUIDE&dl=GUIDE&CFID=52920346&CFTOKEN=27345131](http://portal.acm.org/ft_gateway.cfm?id=1183608&type=pdf&coll=GUIDE&dl=GUIDE&CFID=52920346&CFTOKEN=27345131).
- [32] Ochoa Xavier y Duval Erik (2008), Quantitative Analysis of Learning Object Repositories" , Consultado en octubre de 2009 en <http://bobbebaggio.com/Presentations/PADLA111208/ROL-Repositories.pdf>.
- [33] Özgür Arzucan y Güngör Tunga. (2004), Supervised and Unsupervised Machine Learning Techniques for Text Document Categorization, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.123.1804>.
- [34] Parakhin Mikhail y Haluptzok Patrick. (2009), "Finding The Most Probable Ranking of Objects with Probabilistic Pairwise Preferences", Consultado en octubre de 2009 en <http://www.cvc.uab.es/icdar2009/papers/3725a616.pdf>.
- [35] Peters Gabriele, Kruger Norbert y Von Der Malsburg Christoph (1997), Learning object representations by clustering banana wavelet responses, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.50.4104>.
- [36] Qin Tao, Liu Tie-Yan, Zhang Xu-Dong, Wang De-Sheng, Xiong Wen-Ying y Li Hang. (2008), "Learning to Rank Relational Objects and Its Application to Web Search" , Consultado en octubre de 2009 en <http://www2008.org/papers/pdf/p407-qinA.pdf>.
- [37] Repp Stephan, Linckels Serge y Meinel Christoph. (2007), "Towards to an automatic semantic annotation for multimedia learning objects", Consultado en marzo de 2009 en [http://portal.acm.org/ft\\_gateway.cfm?id=1290149&type=pdf&coll=GUIDE&dl=GUIDE&CFID=52920346&CFTOKEN=27345131](http://portal.acm.org/ft_gateway.cfm?id=1290149&type=pdf&coll=GUIDE&dl=GUIDE&CFID=52920346&CFTOKEN=27345131).
- [38] Rosales Rómer y Fung Glenn. (2006), "Learning sparse metrics via linear programming", Consultado en octubre de 2009 en [http://portal.acm.org/ft\\_gateway.cfm?id=1150444&type=pdf&coll=ACM&dl=ACM&CFID=52920346&CFTOKEN=27345131](http://portal.acm.org/ft_gateway.cfm?id=1150444&type=pdf&coll=ACM&dl=ACM&CFID=52920346&CFTOKEN=27345131).

- [39] Seidl Thomas, Muller Emmanuel, Assent Ira y Steinhausen Uwe (2009), "Outlier Detection and Ranking Based on Subspace Clustering", Consultado en octubre de 2009 en <http://drops.dagstuhl.de/opus/volltexte/2009/1934/pdf/08421.SeidlThomas.Paper.1934.pdf>.
- [40] Stephenson Jhon y Sangrà Albert, Fundamentos del diseño técnicopedagógico, P06/M1103/01178: Universitat Oberta de Catalunya (UOC).
- [41] Strehl Alexander y Ghosh Joydeep. (2003), Relationship-based clustering and visualization for high-dimensional data mining, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.19.7967>.
- [42] Shaban Khaled, Basir Otman y Kamel Mohamed (2003), Learning Objects Clustering based on Semantic Understanding of Text, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.428>.
- [43] Vidal Christian L., Segura Alejandra A. y Prieto Manuel E. (2008), "Calidad en objetos de aprendizaje" \_ Consultado en octubre de 2009 en [http://www.web.upsa.es/spdece08/contribuciones/139\\_CalidadEnObjetosDeAprendizajeTypeInstSpringerFinalVidalSeguraPrieto99.pdf](http://www.web.upsa.es/spdece08/contribuciones/139_CalidadEnObjetosDeAprendizajeTypeInstSpringerFinalVidalSeguraPrieto99.pdf).
- [44] Wiley David. (2000), "Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy". The Instructional Use of Learning Objects: Online Versión, Consultado en febrero de 2009 <http://reusability.org/read/chapters/wiley.doc>.
- [45] Woznica Adam, Kalousis Alexandros y Hilario Melanie (2007), "Learning to combine distances for complex representations" , Consultado en octubre de 2009 en [http://portal.acm.org/ft\\_gateway.cfm?id=1273626&type=pdf&coll=GUIDE&dl=GUIDE&CFID=52920346&CFTOKEN=27345131](http://portal.acm.org/ft_gateway.cfm?id=1273626&type=pdf&coll=GUIDE&dl=GUIDE&CFID=52920346&CFTOKEN=27345131).
- [46] Yeh Jen-Yuan, Lin Jung-Yi, Ke Hao-Ren y Yang Wei-Pang (2007), "Learning to Rank for Information Retrieval Using Genetic Programming", Consultado en octubre de 2009 en <http://jenyuan.yeh.googlepages.com/jyyeh-LR4IR07.pdf>.
- [47] Zhao Ying y Karypis George (2002), Soft Clustering Criterion Functions for Partitional Document Clustering, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.5.8340>.
- [48] Zhao Ying y Karypis George. (2001), Criterion functions for document clustering: Experiments and analysis, Consultado en marzo de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.16.6872>.