PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

IMPACTO DE METADATA BASADA EN CURRÍCULUM EN LA EFICIENCIA DE SISTEMAS DE RECOMENDACIÓN EN EDUCACIÓN

MONTSERRAT PERALTA COSTOYA

Tesis para optar al grado de
Magister en Ciencias de la Ingeniería

Profesor Supervisor:
ROSA ALARCÓN CHOQUE

Santiago de Chile, Agosto 2013

© MMXIII, MONTSERRAT PERALTA COSTOYA

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

IMPACTO DE METADATA BASADA EN CURRÍCULUM EN LA EFICIENCIA DE SISTEMAS DE RECOMENDACIÓN EN EDUCACIÓN

MONTSERRAT PERALTA COSTOYA

Miembros del Comité:
ROSA ALARCÓN CHOQUE
KARIM PICHARA
HERNÁN ASTUDILLO
JUAN DE DIOS ORTÚZAR

Tesis para optar al grado de
Magister en Ciencias de la Ingeniería

Santiago de Chile, Agosto 2013

© MMXIII, MONTSERRAT PERALTA COSTOYA

A mi familia y todos aquellos que
me han acompañado en esta etapa

AGRADECIMIENTOS

Me gustaría agradecer a todos aquellos que me han apoyado y acompañado en esta etapa de aprendizaje.

En primer lugar, quisiera agradecer a Cristóbal Undurraga, por sus consejos, conocimientos y constante apoyo en el desarrollo de esta tésis. Además me gustaría dar las gracias al profesor Karim Pichara, por su asesoría en temas de aprendizaje de máquina y modelos probabilísticos. A mi profesora guía, Rosa Alarcón, por su constante asesoría, apoyo y enseñanza a lo largo de este proceso. Y finalmente, dar las gracias a mi familia, por siempre confiar en mí y en lo que puedo lograr.

INDICE GENERAL

AGRADECIMIENTOS iv

INDICE DE FIGURAS vii

INDICE DE TABLAS viii

RESUMEN ix

ABSTRACT x

1. INTRODUCCION 1

1.1. Motivación 3

1.2. Defínción del Problema 5

1.2.1. Hipótesis 5

1.2.2. Objetivos 5

1.3. Métodos Existentes 6

1.3.1. Sistemas de Recomendación 6

1.3.2. Descripción de los Recursos Digitales de Aprendizaje 8

1.4. Desventajas de los Enfoques Existentes 9

1.5. Contribuciones Originales 10

1.6. Organización de la Tesis/Documento 11

2. MARCO TEORICO 12

2.1. Sistemas de Recomendación 12

2.1.1. Recomendaciones por Contenido 12

2.1.2. Recomendaciones Colaborativas 13

2.1.3. Sistemas de Recomendación Híbridos 15

2.1.4. Preferencias del Usuario 17

2.1.5. Evaluación 17
2.2. Sistemas de Recomendación en Educación 18
2.2.1. Particularidades Escenario Educativo 18
2.2.2. Ejemplos de Sistemas de Recomendación en Educación 20
2.3. Tipos de Información para Describir Recursos Digitales de Aprendizaje . . 23
3. METODOLOGIA 26
4. ANALISIS PRELIMINAR 29
4.1. Sistema de Recomendación Adecuado para el Escenario Educativo del Portal EducarChile 29
5. RESULTADOS 35
5.1. Evaluación Sistema de Recomendación Educacional Colaborativo, por Contenido e Híbrido 35
5.2. Encontrando la Metadata en nuestro Conjunto de Datos Educacional . . . 37
6. CONCLUSIONES Y TRABAJO FUTURO 42
6.1. Revisión de los Resultados y Comentarios Generales 42
6.2. Temas de Investigación Futura 43
BIBLIOGRAFIA 44

INDICE DE FIGURAS

4.1 Modelo Gráfico de Sistemas de Recomendación Probabilísticos Comparados en esta Investigación 33

5.1 Comparación Sistemas de Recomendación Probabilísticos 36

5.2 Comparación Tipos de Información para Describir Recursos Digitales de Aprendizaje 39

INDICE DE TABLAS

5.1 Comparación Sistemas de Recomendación que Utilizan Estimación de Ratings 36

5.2 Ejemplos de Grupos de Información Generados no Considerados por Metadata Curricular 41

RESUMEN

Los profesores se encuentran a diario con la tarea de recopilar material para preparar sus lecciones. Muchas veces el proceso de crear material nuevo es difícil y el tiempo disponible es escaso. En internet existen repositorios de recursos digitales de aprendizaje especializados; el problema es que, debido a la gran cantidad y variedad de recursos disponibles, la tarea de encontrar un recurso de calidad y adecuado al propósito educativo se vuelve compleja. Los sistemas de recomendación constituyen una herramienta adecuada para apoyar este proceso de búsqueda y descubrimiento; técnicas de filtrado colaborativo aprovechan el comportamiento de la comunidad, enfoques basados en contenido se enfocan en la descripción de los recursos para encontrar ítems similares a recomendar, mientras que los enfoques híbridos combinan filtrado colaborativo y por contenido. En esta investigación se comparó varios sistemas de recomendación en un escenario educativo real, correspondiente al portal educacional Educarchile. Se encontró que el sistema de recomendación híbrido *Collaborative Topic Regression* (CTR), que saca provecho de las características de los usuarios y de la descripción de los recursos, es el que mejor se adecúa al escenario educacional estudiado. Además, se pudo concluir que describir los recursos digitales de aprendizaje a través de metadata curricular (nivel, asignatura, etc.) mejora la precisión del sistema de recomendación, mientras que el texto libre u otro tipo de metadata impacta negativamente en los resultados.

Palabras Claves: Sistemas de Recomendación, Tecnología en Educación, Modelos Probabilísticos, Metadata,

ABSTRACT

Finding the proper material to enrich or create educational content is an everyday task for most teachers. Educational content sources are spread over the Web and provided in specialized repositories. Content abundance and variety makes searching a complex task. Recommender systems can assist teachers in finding the proper content; collaborative filtering techniques exploits community's behavior typically captured in a dataset, content-based approaches focus on content descriptions to find similar items to recommend whereas hybrid approaches mix collaborative filtering and content-based recommendations. In this study we compared different recommendation techniques in a real educational scenario, corresponding to the educational portal EducarChile. We found that the Collaborative Topic Regression (CTR) hybrid recommendation system, that takes advantage of the characteristics of the users and of the description of resources, appears to be best suited for the educational setting of EducarChile. Furthermore, we conclude that describing digital learning resources through curriculum metadata (level, subject, etc.) had a positive impact on recommender's accuracy whereas free-text or other metadata negatively impacted the results.

Keywords: Recommender Systems, Technology Enhanced Learning, Probabilistic Models, Metadata.

1. INTRODUCCION

A diario los profesores se encuentran con la tarea de recopilar material para preparar sus lecciones. Muchas veces el proceso de crear un nuevo material es difícil y el tiempo disponible para hacerlo es escaso. Por esta razón los profesores antes de crear algo nuevo buscan material de calidad existente que se adecúe a sus necesidades y que puedan reutilizar para enriquecer sus lecciones. Se ha mostrado que el 92% de los educadores ha buscado en internet recursos digitales de aprendizaje (Petrides, Nguyen, Kargliani, & Jimes, 2008). Dentro de las razones de su búsqueda están: el deseo de incluir nuevo material en sus cursos, mejorar sus habilidades de enseñanza, mejorar su conocimiento y conectarse con colegas que tienen intereses de enseñanza similares. Para profesores y también para estudiantes, la proliferación de recursos digitales de aprendizaje ha significado acceso centralizado a materiales que satisfacen necesidades únicas de enseñanza-aprendizaje, y la posibilidad de colaborar con sus pares para crear, usar y reutilizar este tipo de recursos, incluso de manera global.

Dentro de la utilización y reutilización de recursos digitales podemos encontrar cuatro fases (Clements & Pawlowski, 2012):

(i) Fase de búsqueda: ¿Dónde y cómo pueden los profesores encontrar recursos adecuados para su uso?

(ii) Fase de evaluación: ¿Son los recursos encontrados adecuados para el uso que quiere darle el profesor? ¿Son adaptables para el contexto que se busca? Aquí se toma en cuenta también la calidad del recurso, ¿Es el recurso recomendado por alguna institución de importancia o por algún experto? ¿Qué ha opinado la comunidad acerca del recurso?

(iii) Fase de adaptación: Modificar el recurso educacional al contexto, combinarlo con otros recursos.

(iv) Fase de compartir: Compartir el nuevo recurso ajustado de regreso a la comunidad.

Hay evidencia que los profesores están utilizando y reutilizando recursos digitales, siguiendo las fases anteriores, a distintos niveles (Pegler, 2012), desde búsquedas de material en un solo repositorio, búsquedas a través de varios repositorios (incluso de otros países) (Vuorikari & Koper, 2009b), hasta contribuciones creando nuevos materiales o adaptando existentes. El 89% de los profesores encuentra recursos buscando en tópicos y asuntos. El 82% encuentra recursos basados en recomendaciones de colegas. El 71% encuentra recursos basados en recomendaciones de amigos personales. El 56% buscan por recursos bien rankeados. El 58% buscan por recursos que provengan de una organización con buena reputación (Clements & Pawlowski, 2012). Con lo anterior no solo podemos ver que los usuarios están buscando recursos sino que la confianza es un instrumento clave de calidad. Confían en que un recurso será mejor si está bien evaluado por alguien que conocen o por una institución de prestigio. Es así que la calidad de lo que el usuario encuentre en su búsqueda es un aspecto muy importante.

En internet existen repositorios especializados para buscar este tipo de recursos. Dentro de los más conocidos a nivel mundial se encuentran MACE (Wolpers, Memmel, & Giretti, 2009)[1], Mendeley[2], MERLOT (Sicilia, García-Barriocanal, Sánchez-Alonso, & Cechinel, 2010)[3] y OER Commons[4]. En Chile, particularmente, tenemos Educarchile[5] y PortalEducativo[6], entre otros.

Todos estos repositorios tienen una enorme cantidad de recursos digitales disponibles, variando en su origen, calidad, propósito, etc. Esto da la opción de poder acceder a material

[1]htpp://www.mace-project.eu
[2]http://www.teleurope.eu/pg/pages/view/50630/
[3]http:// www.merlot.org
[4]http://www.oercommons.org
[5]http://www.educarchile.cl
[6]http://www.portaleducativo.net/

de calidad de forma centralizada, permitiendo a distintos profesores compartir con sus pares, recomendar recursos de interés y encontrar material novedoso y de calidad para preparar sus lecciones.

1.1. Motivación

Los repositorios de recursos digitales de aprendizaje representan para los profesores una forma centralizada de acceder a material para utilizar en sus lecciones. A pesar de la gran ventaja que significan estos repositorios, muchas veces la tarea de encontrar un recurso que se adapte al propósito educativo y que sea de la calidad deseada se hace muy difícil, debido a la gran cantidad de opciones donde escoger, y la diversidad de material disponible. Además en ocasiones las tecnologías utilizadas para realizar las búsquedas de material no son las apropiadas, o las categorías de información que se utilizan para describir a los recursos no son las correctas. Por ejemplo, al buscar utilizando la palabra "célula", se pueden obtener resultados relacionados con biología básica (ej. descripción de la célula, mitosis, meiosis, etc.) y otros más avanzados (ej. papers relacionados con nuevos descubrimientos en medicina referidos a procesos celulares), siendo estos últimos resultados inadecuados para un profesor de enseñanza básica. Lo anterior lleva a que los usuarios no encuentren material relevante, siendo éste uno de los principales desincentivos a la hora de buscar y reutilizar recursos digitales (Petrides et al., 2008).

Dentro de las herramientas que ayudan a mejorar el descubrimiento de recursos de calidad se encuentran los sistemas de recomendación (Manouselis, Vuorikari, & Van Assche, 2007). Estos sistemas utilizan las preferencias o características de los usuarios (sistemas de recomendación colaborativos), o el contenido de los recursos visitados previamente (sistemas de recomendación basados en contenido), o incluso una combinación de las anteriores (sistemas de recomendación híbridos), para proponer un nuevo recurso que podría ser interesante para el usuario. Con esto, los recursos

propuestos al usuario son más adecuados a sus preferencias, ya que toman en consideración información previa en un proceso más personalizado de búsqueda. Hay que tener en cuenta que existen muchos sistemas de recomendación distintos (Adomavicius & Tuzhilin, 2005) y el dominio educacional es un dominio rico contextualmente, donde no sólo hay distintos tipos de recursos y actividades a recomendar sino que además se busca lograr una meta educacional (Santos & Boticario, 2010). Es por lo anterior que no es cualquier sistema de recomendación el que debe utilizarse en un repositorio de recursos digitales de aprendizaje. Hay que hacer una evaluación para encontrar el que mejor se adecúe al escenario, a los datos y al público objetivo (Drachsler et al., 2010). Además, se debe encontrar la correcta parametrización del sistema que se ajuste al repositorio específico que se quiera utilizar (Manouselis, Drachsler, Vuorikari, Hummel, & Koper, 2011).

Por otra parte, la representación de los recursos digitales es muy importante. Ésta permite facilitar su búsqueda y descubrimiento. Los recursos digitales pueden ser descritos por su contenido (palabras que forman un archivo de texto, descripción de una fotografía, resumen de un sitio educativo, etc.) o pueden ser clasificados utilizando metadata, la cual puede ir desde una estructura bien definida (ontología) (Nikolopoulos, Kalou, Pierrakeas, & Kameas, 2012) hasta *tags* libres que puedan incluir los usuarios permitiendo clasificar los recursos (folksonomías) (Vuorikari & Põldoja, 2009). La metadata considerada puede ser relacionada con el currículum al cual pertenece el recurso, su nivel, asignatura, etc. (que llamaremos metadata curricular) o asociada a otras áreas de conocimiento de los profesores como objetivo pedagógico, público objetivo, método de enseñanza, tecnología utilizada, etc. (que llamaremos metadata pedagógico tecnológica) (Mishra & Koehler, 2006). Actualmente no hay un consenso sobre cuál es la mejor para describir recursos digitales (Hummel et al., 2007).

1.2. Defincián del Problema

En este investigación nos interesa, en primer lugar, encontrar qué características deben tener las tecnologías de búsqueda y recomendación de material y en segundo, cuáles deben ser las categorías de información que permitan mejorar el descubrimiento y utilización de material de calidad por parte de los usuarios (profesores) de repositorios de recursos digitales de aprendizaje.

1.2.1. Hipótesis

Considerando que las bases de datos de los repositorios de recursos digitales de aprendizaje son de carácter escaso (*sparse*), esto es, las visitas de los usuarios no cubren todo el universo de recursos disponibles y que queremos novedad en los recursos que se recomienden a los profesores, la primera hipótesis de este trabajo es que en el área de educación un sistema de recomendación híbrido debiese ser el que genera mejores resultados, en cuanto a precisión y *recall*, en comparación con sistemas de recomendación colaborativos y por contenido.

Por otra parte, considerando que el contexto de educación básica y media está fuertemente regulado por currículum fijos en Chile y Latinoamérica y que, por lo mismo, el descubrimiento de información estaría fuertemente influenciado por este contexto, la segunda hipótesis de este trabajo es que anotar los recursos educativos con metadata curricular y no solo utilizar su contenido propio o considerar otros metadatos, mejora el *recall* y precisión de los sistemas de recomendación.

1.2.2. Objetivos

Nuestro primer objetivo es evaluar y comparar las principales estrategias de sistemas de recomendación y determinar cuál de ellas genera mejores resultados, en cuanto a *recall*

y precisión, en un repositorio de recursos digitales de aprendizaje real. Lo que permite ayudar a la búsqueda y descubrimiento de nuevos recursos por parte de los profesores.

El segundo objetivo de esta investigación es evaluar qué tipo de información influye en las métricas de *recall* y precisión generadas por el mejor sistema de recomendación encontrado para el escenario educativo. Esto debiera permitir determinar cómo debe describirse los recursos digitales.

1.3. Métodos Existentes

A continuación describiremos los métodos existentes de sistemas de recomendación y las formas utilizadas para describir lo recursos a recomendar.

1.3.1. Sistemas de Recomendación

Los sistemas de recomendación se han utilizado desde ya varios años para ayudar a mejorar la búsqueda de recursos en internet. Hay varios enfoques según cómo se realice la recomendación de un nuevo recurso:

(i) Sistemas de recomendación por contenido (Lops, Gemmis, & Semeraro, 2011): Al usuario se le recomienda recursos similares a aquellos que ha preferido en el pasado.

(ii) Sistemas de recomendación colaborativos (Schafer, Frankowski, Herlocker, & Sen, 2007): Al usuario se le recomienda recursos que personas con gustos similares (inferidos) han preferido en el pasado.

(iii) Sistemas híbridos (Burke, 2007): Son métodos que combinan los dos tipos de métodos anteriores.

Para realizar estas recomendaciones las preferencias de los usuarios se pueden describir de dos maneras: *ratings*[7] o evaluaciones explícitas, donde se le pide al usuario que evalúe un recurso en una cierta escala, o *ratings* implícitos, donde, a través de acciones que realiza el usuario, como visitas a *urls* de recursos, descarga de éstos, etc., se infiere si un recurso le parece interesante.

Los algoritmos utilizados por los sistemas de recomendación se pueden clasificar en dos tipos: basados en memoria y basados en modelos. Los primeros utilizan heurísticas que hacen predicciones de *ratings* basados en la colección entera de recursos previamente vistos por el usuario, computando todo en memoria. Los algoritmos basados en modelos, en cambio, usan la colección de *ratings* de los recursos previamente vistos por los usuarios, para formular un modelo, el cual es luego usado para calcular la probabilidad de que un nuevo recurso sea interesante para el usuario.

A lo largo de los años ha habido muchas implementaciones de los distintos enfoques descritos anteriormente, principalmente en rubros como recomendación de productos, libros, películas, etc. Además de esto, y más recientemente, los sistemas de recomendación se han empezado a utilizar en ambientes distintos, como Educación.

En Educación el escenario al momento de recomendar un recurso es un poco distinto (Manouselis et al., 2011), debido a que los usuarios buscan lograr metas de aprendizaje, i.e tienen un propósito educativo. Además hay que considerar que puede haber distintos tipos de escenario: escenarios formales (participan expertos en un currículum fijo) o informales (estudiantes realizan un proceso de auto-aprendizaje a través de internet). Así mismo, hay que tener en cuenta que el público objetivo puede ser muy variado (profesores, académicos, estudiantes, etc.). Por todo lo anterior, los sistemas de recomendación aplicados a

[7]Los *ratings* son evaluaciones que los usuarios hacen a los recursos. Generalmente van en una escala de 1 a 5, indicando con 5 que el recurso les pareció interesante y con 1 totalmente lo contrario

Educación no son iguales a los aplicados a otros rubros ya que deben considerar todos los aspectos previamente descritos.

1.3.2. Descripción de los Recursos Digitales de Aprendizaje

El tipo de información que describe a los recursos digitales en educación influye en el proceso de búsqueda y descubrimiento de éstos (Vuorikari & Koper, 2009a). Hay varias formas de describir los recursos. La más fácil, y muchas veces la más usada, es el contenido propio del recurso, como el texto que compone un archivo, descripción de una fotografía, resumen de un sitio web, etc. Además, de esta forma, un recurso digital de aprendizaje se puede describir a través de metadata curricular, identificando principalmente a qué nivel educacional corresponde, que unidad está evaluando, etc (Becerra, Astudillo, & Mendoza, 2012). Por otra parte, la metadata que describe a un recurso digital de aprendizaje puede ser de tipo pedagógico-tecnológica (Mishra & Koehler, 2006), es decir, centrada en indicar cuál es el objetivo pedagógico del recurso, cuál es la audiencia objetivo, cuáles son sus características técnicas, etc.

Dentro de la metadata estructurada el estándar más conocido es *Learning Object Metadata* LOM (Technology & Committee, 2002) que especifica la sintaxis y semántica de la metadata de los objetos de aprendizaje. Este estándar se centra en el mínimo conjunto de atributos necesarios para permitir que estos objetos sean manejados, encontrados y evaluados. Este tipo de metadata permite desligarse del contenido de los recursos e incluso puede generarse automáticamente (Riley, Miller, Soh, Samal, & Nugent, 2009). Dentro de la forma más libre de describir los recursos, además del contenido propio, se encuentran los *tags* (etiquetas), palabras que asignan libremente los usuarios para clasificar los recursos. Hay muchos sistemas de recomendación que se basan en *tags*, ya que éstos traen beneficios como flexibilidad, rápida adopción y fácil usabilidad (Krestel & Fankhauser, 2012)

1.4. Desventajas de los Enfoques Existentes

Dentro de los sistemas de recomendación cada enfoque tiene sus desventajas. Los sistemas de recomendación por contenido se enfrentan a problemas de semántica y sinonimia, además sufren de sobre-especialización ya que sólo se les recomienda viendo lo que el mismo usuario ha visto en el pasado, por lo que no hay novedad. Además se necesita información considerable de recursos vistos previamente para hacer una nueva recomendación.

Los sistemas de recomendación colaborativos también se encuentran con problemas cuando llega un nuevo usuario pues se debe tener una historia de las preferencias del usuario para recomendarle algo nuevo; lo mismo ocurre cuando ingresa un nuevo ítem, el cual debe ser evaluado por varios usuarios antes de considerarse para recomendar.

Por su parte, los sistemas híbridos buscan sopesar estas desventajas, considerando lo mejor de los sistemas colaborativos y por contenido, pero igual caen en desventajas principalmente por su complejidad al momento de construirlos.

Además de lo anterior, una gran desventaja que tienen los sistemas de recomendación en el ámbito educacional es que no se toma en consideración las peculiaridades específicas de este escenario y se les aplica tal como si se comportaran igual que al recomendar cualquier producto. Por otra parte, hay muy pocos repositorios de recursos digitales en el ambiente de educación que sean públicos, por lo que en general los sistemas de recomendación no se prueban en ambientes reales, (Verbert et al., 2011) y se terminan evaluando en otros escenarios, dando resultados teóricos pero que podrían variar en un repositorio educativo real.

En cuanto al tipo de información utilizada para describir los recursos digitales de aprendizaje, encontramos desventajas al utilizar metadata estructurada (Di Nitto, Mainetti, Monga, Sbattella, & Tedesco, 2006): muchas veces el significado exacto de cierta metada

no está claramente especificada, alguna no es procesable por máquina, e incluso algunas características de los recursos no pueden ser expresadas. Los *tags*, por su parte, también tienen desventajas ya que sus significados tienen que ser evaluados e inferidos en el contexto de otros *tags* e información del usuario y algunos *tags* incluso pueden tener significado sólo para el usuario que los creó. Lo anterior pasa también con el contenido propio, además de problemas con la sinonimia y semántica, ya que las palabras se ven sólo como conjunto de letras, por lo que no se toma en cuenta que una palabra puede tener sinónimos u otras palabras semánticamente parecidas que debieran ser consideradas como iguales.

1.5. Contribuciones Originales

Los métodos de análisis de esta investigación se han aplicado a una base de datos real representante del escenario educativo en Chile, donde los usuarios han realizado actividades durante 3 años, de forma natural (no controlada, sin instrucciones). Esta investigación se centra en un repositorio educarivo real y destinado a la educación básica y media, a diferencia de la mayoría de los estudios de este tipo que se centran en repositorios no educacionales o destinados a la educación universitaria y de aprendizaje continuo.

A diferencia de otros enfoques, en esta investigación se están analizando sistemas de recomendación sofisticados, que en lugar de basarse en cómputos en memoria y fuerza bruta, utilizan un modelo probabilístico sobre los datos. Los resultados de precisión y *recall* son mucho mejores y se puede dar una interpretación de lo que está ocurriendo con la clasificación de recursos para recomendar.

Además de lo anterior, a diferencia de estudios relacionados, al enfocarnos en el sistema educativo básico y medio partimos de una estructura de metadatos que refleja este contexto educativo. Por lo tanto, no seguimos un enfoque de publicador de contenido (ej,

un libro), tal como la metada LOM, y vamos más allá que manejar la información como texto plano o a través de *tags* libres.

Presentamos una comparación entre tres tipos de información utilizada para describir recursos digitales de aprendizaje, algo no realizado por otros estudios. Además, nuestros resultados demuestran una diferencia significativa entre estos tres enfoques (contenido, metadata curricular, metadata pedagógico-tecnológica), lo que confirma nuestra hipótesis de que, dado el contexto educativo en el que estamos, describir los recursos a través de metadata curricular ayuda a mejorar el proceso de búsqueda y descubrimiento utilizando sistemas de recomendación.

1.6. Organización de la Tesis/Documento

El trabajo se organiza de la siguiente manera: en el capítulo 2 presentamos los distintos sistemas de recomendación existentes y cuáles deben ser las consideraciones a tomar en escenarios educativos. Hacemos una revisión de los distintos sitemas de recomendación actualmente aplicados en educación y los distintos tipos de información que se utilizan para describir los recursos digitales; en el capítulo 3 vemos la metodología utilizada y las características con que se desarrolló la experimentación; en el capítulo 4 hacemos un análisis teórico para determinar cuáles deberían ser los sistemas de recomendación que mejor se adecúen al contexto educacional en el que nos encontramos; en el capítulo 5 mostramos los resultados obtenidos tanto en la comparación entre sistemas de recomendación, como en la influencia que tienen las distintas formas de describir a los recursos digitales de aprendizaje y en el capítulo 6 presentamos las conclusiones y análisis de los posibles trabajos a futuro.

2. MARCO TEORICO

En esta sección analizaremos los distintos tipos de sistemas de recomendación existentes y cuáles deben ser las consideraciones a tomar para aplicarlos a escenarios educativos. Revisaremos los distintos sistemas de recomendación actualmente aplicados en educación y los distintos tipos de información que se utilizan para representar los recursos digitales a ser recomendados.

2.1. Sistemas de Recomendación

Los sistemas de recomendación utilizan características conocidas de preferencias de los usuarios para recomendar un nuevo recurso, previamente no visto por el usuario, que se ajuste a esas preferencias (Schafer et al., 2007). Existen varias formas de abordar esta tarea, las cuales se pueden clasificar en tres enfoques distintos: recomendaciones por contenido, recomendaciones colaborativas y sistemas híbridos (Adomavicius & Tuzhilin, 2005), los cuales de describen a continuación

2.1.1. Recomendaciones por Contenido

Al usuario se le recomienda recursos similares a aquellos que ha preferido en el pasado. Uno de los métodos más utilizado es *Term Frequency-Inverse Document Frequency* TF-IDF (Ramos, 2003) o alguna heurística de calificación en términos de vectores. En estos algoritmos se le asigna un peso a cada palabra que describe el contenido de los recursos y se ve a estos últimos como vectores de palabras, midiendo así la similitud entre recursos como similitud entre vectores. También se han utilizado Clasificadores Bayesianos (Billsus & Pazzani, 1998), y otras técnicas de aprendizaje de máquina como Clustering (Shepitsen, Gemmell, Mobasher, & Burke, 2008), Árboles de Decisión (Cho, Kim, & Kim, 2002) y Redes Neuronales Artificiales (Rumelhart, Hinton, & Williams, 2002). Estos últimos ya no se basan en una heurística, o medida de similitud entre vectores, sino en identificar un

modelo aprendido de la información subyacente y luego utilizar ese modelo para calcular la probabilidad de que un nuevo recurso sea interesante para el usuario (Ono, Kurokawa, Motomura, & Asoh, 2007; Cho et al., 2002).

Otros métodos consideran el uso de ontologías (Middleton, Shadbolt, & De Roure, 2004; Aciar, Zhang, Simoff, & Debenham, 2007), o utilizar la semántica para enriquecer el contenido considerado de los recursos (Gabrilovich & Markovitch, 2006, 2007). Finalmente, se ha considerado también el uso de Modelos Temáticos (D. Blei, Carin, & Dunson, 2012) que son modelos probabilísticos que permiten descubrir la estructura semántica de una colección de documentos sobre la base de un análisis bayesiano jerárquico del texto original. Por ejemplo, *Latent Dirichlet Allocation* LDA (D. M. Blei, Ng, & Jordan, 2003) es un modelo bayesiano jerárquico de tres niveles, en donde cada ítem de una colección es modelado como una mixtura finita sobre un conjunto subyacente de tópicos.

Ventajas: (Lops et al., 2011) Independencia de los demás usuarios, sólo se tiene en consideración lo que el usuario ha visto previamente. Transparencia, se puede dar una explicación de cómo funciona el sistema de recomendación. Capacidad de recomendar un recurso nuevo que aún no ha sido visto por ninguna otra persona.

Desventajas: Sólo toma en consideración palabras, por lo que no se puede identificar entre un artículo mal escrito de uno bien escrito, además de problemas como polisemia y sinonimia. Sobre-especialización, sólo se le recomienda parecido a lo que el usuario ha visto en el pasado, no hay novedad. Problema del nuevo usuario: se necesita un número considerable de recursos previamente vistos para hacer una recomendación

2.1.2. Recomendaciones Colaborativas

Al usuario se le recomienda recursos que personas con gustos similares han preferido en el pasado.

Estos sistemas se pueden dividir en basados en memoria y basados en modelos (Adomavicius & Tuzhilin, 2005), según el algoritmo que utilicen. Los primeros son esencialmente heurísticas que hacen predicciones de *ratings* basados en la colección entera de recursos evaluados previamente por el usuario (métodos más utilizados: Correlación de Pearson, Basados en Coseno (Schafer et al., 2007)). Los últimos, en contraste, usan la colección de *ratings* (de los recursos previamente vistos por los usuarios) para aprender un modelo, el cual es luego usado para hacer predicciones de nuevos *ratings* (métodos más usados: *Clustering* y Redes Bayesianas).

De los sistemas basados en memoria los más conocidos son *user-based nearest neighbor* (Lops et al., 2011) e *ítem-based nearest neighbor* (Sarwar et al., 2001). Uno de los problemas de *user-based nearest neighbor*, utilizando correlación de Pearson, es que sólo mide la similitud entre dos usuarios y no considera lo que opina la mayoría. Por ejemplo, un recurso visto por más usuarios (recurso popular) debería ser menos relevante cuando se compara los recursos que tienen en común dos usuarios. Así, los recursos menos populares (controversiales) aportan más a la similitud entre esos dos usuarios (Schafer et al., 2007). Además, el cómputo de estos algoritmos es caro dado que cada usuario hay que compararlo con todos los demás. Se pueden usar pre-computaciones y *Clustering* para superar esto, pero aún así el algoritmo requiere mucha memoria (Adomavicius & Tuzhilin, 2005). Los *ítem-based nearest neighbor* son la transpuesta del anterior y hay evidencia que dice que son más acertados que los centrados en usuarios (Schafer et al., 2007).

De los sistemas basados en modelos unos de los más utilizados son Clasificadores de Bayes (Chen & George, 1999; Condliff, Madigan, Lewis, & Posse, 1999). Una limitación de los Clasificadores de Bayes es que un usuario puede sólo ser clasificado en una clase, sin considerar que algunas aplicaciones de recomendación se pueden beneficiar de la habilidad de clasificar a varios usuarios en varias clases a la vez.

Hay también métodos que mezclan algoritmos basados en memoria y basados en modelos como (Pennock, Horvitz, Lawrence, & Giles, 2000) que evalúa el método llamado *Personality Diagnosis*. Dadas las preferencias de un usuario por algunos ítems se calcula la probabilidad que ese usuario sea del mismo "tipo de personalidad" que otros, y luego la probabilidad de que a ese usuario le gusten los nuevos ítems. Demuestra ser mejor que métodos basados en memoria como Correlación y Similitud de vectores, y métodos basados en modelos como Clasificadores y Redes de Bayes.

Ventajas: El contenido no es importante, sólo lo que otros usuarios han visto, así que no hay limitación desde este aspecto.

Desventajas: Un nuevo usuario debe tener suficientes recursos previamente vistos. Un nuevo recurso debe ser evaluado por varios usuarios antes de ser considerado. En general hay mucha escasez de datos pero se puede manejar con filtrado dmográfico, dando primero recomendaciones según el segmento demográfico al que pertenezca cada usuario dadas sus características.

2.1.3. Sistemas de Recomendación Híbridos

Son sistemas que combinan los dos tipos de métodos anteriores. Así, varios sistemas de recomendación se utilizan como componentes de un sistema más grande que entrega finalmente los recursos a recomendar. Según como se integren los componentes de recomendación se puede identificar siete tipos de híbridos (Burke, 2007):

(i) Ponderado: La puntuación de distintos componentes de recomendación se combinan numéricamente. Es el más simple. Se usa una fórmula lineal para combinarlos.

(ii) Conmutación: El sistema elige entre componentes de recomendación y aplica el elegido. Según la situación elige uno u otro.

(iii) Mixto: Recomendaciones de diferentes sistemas de recomendación son presentadas juntas. Se muestran una al lado de la otra en una lista combinada.

(iv) Combinación de características: Características derivadas de distintas fuentes de conocimiento son combinadas y derivan en un solo algoritmo de recomendación. Toma prestada la lógica de recomendación de otra técnica, más que emplear un componente separado que lo implemente.

(v) Incremento de características: Una técnica de recomendación es usada para computar una característica o conjunto de características que luego son parte del input de la siguiente técnica. Se usa para añadir una fuente de conocimiento cuando ya hay un componente fuerte de recomendación.

(vi) Cascada: Se da prioridades a los componentes de recomendación y los más débiles no pueden cambiar las recomendaciones de los más fuertes, sólo refinarlas.

(vii) Meta-nivel: Una técnica de recomendación es aplicada y produce alguna clase de modelo, el cual es luego el input usado por la siguiente técnica. Aquí el primero genera la fuente de conocimiento para la segunda, dándole el modelo generado.

En general los sistemas híbridos de recomendación teóricamente debiesen ser los mejores pues complementan las ventajas de los sistemas colaborativos y por contenido, superando las desventajas que cada uno de ellos presenta por separado.

(Yu, Schwaighofer, Tresp, Ma, & Zhang, 2002) es un ejemplo de un sistema de recomendación híbrido entre filtrado colaborativo y filtrado por contenido. Para filtrado colaborativo usa Correlación de Pearson y para filtrado por contenido usa SVM (Support Vector Machines). Los resultados muestran que en cuanto a precisión son más adecuados que filtrado colaborativo o filtrado por contenido por separado.

En este investigación compararemos los distintos tipos de sistemas de recomendación antes mencionados: filtrado colaborativo, filtrado por contenido e híbridos y también los

distintos tipos de algoritmos (basados en memoria, basados en modelos) para encontrar el que mejor se adecúe al escenario particular de este experimento.

2.1.4. Preferencias del Usuario

Los sistemas de recomendación en general utilizan *ratings* entregados por los usuarios para interpretar sus preferencias. Incluso se ha propuesto una mejora en cuanto a utilizar *ratings* multicriterio, donde se evalúen distintos aspectos de los recursos y no sólo se coloque una evaluación general (Manouselis, Vuorikari, & Van Assche, 2010). De todas formas, el pedir al usuario que evalúe un recurso (*rating* explícito) es muy intrusivo. Hay algunos sistemas de recomendación que utilizan las acciones de los usuarios para inferir *ratings* (*rating* implícito), como por ejemplo las visitas que realice a una *url*, siendo así el proceso de evaluación transparente para el usuario, pero claramente no tan certero como pedírselo explícitamente.

En esta investigación utilizaremos un sistema de *ratings* implícitos, analizando los *logs* de *url* de recursos digitales visitadas por los usuario, ya que es importante no agregar complejidad para mejorar el descubrimiento, si no que mantener el proceso lo más simple para el usuario.

2.1.5. Evaluación

La evaluación de los sistemas de recomendación se hace generalmente en términos de cobertura y precisión (Herlocker, Konstan, Terveen, & Riedl, 2004). Cobertura mide el porcentaje de recursos para los cuales un sistema de recomendación es capaz de hacer predicciones. Precisión en términos estadísticos compara los *ratings* estimados contra los *ratings* reales e incluye MAE, RMSE y correlaciones entre predicciones y *ratings*. Precisión determina qué tan bien un sistema de recomendación puede hacer predicciones de ítems de alta relevancia. Ellos incluyen precisión, el porcentaje de realmente “altos” *ratings* dentro de esos que fueron predichos como “altos” por el sistema de recomendación;

recall, el porcentaje de predichos correctamente "altos" *ratings* dentro de todos los *ratings* conocidos por ser "altos". De todas formas hay que tener presente que estas métricas técnicas no capturan ni la usabilidad o calidad de las recomendaciones.

Otra característica deseada en los sistemas de recomendación, más difícil de evaluar, es la novedad de la recomendación, lo mismo con serendipia, la capacidad del sistema de recomendar algo que el usuario no habría encontrado por sí solo.

Nosotros utilizaremos una combinación de métricas técnicas (precisión y *recall*) para evaluar los distintos sistemas de recomendación en el escenario real del experimento. Siendo deseable como trabajo futuro evaluar métricas más cualitativas para poder llegar a conclusiones sobre la calidad, usabilidad y novedad de las recomendaciones.

2.2. Sistemas de Recomendación en Educación

2.2.1. Particularidades Escenario Educativo

En Educación, el escenario al momento de recomendar un recurso es un poco distinto al de recomendar una nueva película o un nuevo objeto que la persona quisiera comprar (Manouselis et al., 2011). Los objetivos de los sistemas de recomendación en educación son diferentes a los de los sistemas de recomendación tradicionales. Entre las tareas que deben apoyar están:

(i) Encontrar recursos interesantes y de buena calidad.

(ii) Encontrar recursos novedosos: esto significa explorar técnicas de recomendación que seleccionen recursos más allá de sus similitudes.

(iii) Encontrar compañeros: explorar atributos para medir la similitud con otras personas.

(iv) Encontrar buenos caminos de aprendizaje: explorar criterios para la construcción y sugerir secuencias alternativas (pero similares).

Además hay que tener en cuenta que existen ambientes formales e informales de aprendizaje (Manouselis et al., 2011). Un ambiente formal de aprendizaje incluye ofertas de aprendizaje de instituciones educacionales con un currículum y es caracterizado como altamente estructurado, que lleva a una acreditación y que involucra a expertos del dominio para asegurar calidad. Esto tradicionalmente ocurre en ambientes dirigidos por profesores con interacciones persona a persona en una forma en vivo y síncrona. Un ambiente informal se refiere a sujetos que no están participando en ningún aprendizaje formal y que son responsables de su propio ritmo y camino de aprendizaje. El proceso de aprendizaje depende en gran medida de preferencias y decisiones individuales y es regularmente auto-dirigida. Los recursos para este tipo de ambiente pueden venir de comunidades de expertos, contexto de trabajo, formación o incluso amigos. En nuestro caso el escenario representa un ambiente formal pues se cuenta con una estructura y expertos que clasifican y analizan los recursos disponibles, los cuales son utilizados por profesores en su proceso de enseñanza.

Por lo anterior, el sistema de recomendación debe también considerar los distintos ambientes y si las recomendaciones se están haciendo a profesores, estudiantes, o un público variado (Bobadilla, Serradilla, & Hernando, 2009). Para contextos basados en profesores hay tareas relacionadas con preparar, dar y evaluar lecciones. Para contextos basados más en estudiantes hay tareas sobre secuencia de actividades a desarrollar o metas para cumplir cierto objetivo de aprendizaje. En nuestro caso nos centraremos solamente en un público compuesto por profesores y estudiantes de pedagogía avanzados, que son los protagonistas en un escenario formal como en el que nos encontramos.

En escenarios informales de aprendizaje los enfoques más utilizados son aquellos basados en lo social (Hummel et al., 2007), como son las folksonomías, *ratings* de usuarios y *tags* libres, mientras que en escenarios formales de aprendizaje se utilizan enfoques más orientados a la información (metadata), como ontologías formales del dominio y modelos de usuarios. Para redes de aprendizaje formales una ecología de metadata híbrida,

es decir, que tenga una parte estructurada pero también permita la libertad de nuevas clasificaciones, debiera ser lo ideal, donde las capas más lentas (ontologías, estructuradas) proveen estabilidad y las capas más rápidas (folksonomías) dirigen al cambio. En nuestro caso queremos ver si realmente el tener metadata híbrida (metadata curricular y metadata pedagógico-tecnológica) es mejor que solamente considerar una de las dos o simplemente utilizar el contenido propio de los recursos (sin metadatos).

Con lo anterior podemos ver que es muy importante no sólo pensar en las características del sistema de recomendación sino en que éstas deben adecuarse a las particularidades que tiene el contexto educativo al que queremos aplicarlo, en nuestro caso, un ambiente formal donde la audiencia son profesores que buscan encontrar material para preparar sus lecciones.

2.2.2. Ejemplos de Sistemas de Recomendación en Educación

Hay muchos ejemplos de sistemas de recomendación en educación (Dron, Siviter, Boyne, & Mitchell, 2000; Farzan & Brusilovsky, 2004; Tang & McCalla, 2004a, 2004b; Manouselis et al., 2007) considerando distintos enfoques y tomando en cuenta lo que hemos detallado en la sección anterior.

Podemos encontrar sistemas de recomendación híbridos (Tang & McCalla, 2004a, 2004b) que mezclan filtrado colaborativo además de filtrado basado en ontologías, utilizando metadata curricular y metadata pedagógico-tecnológica para describirlos. Hay que tener en cuenta que ambos sistemas están limitados a un dominio muy reducido (documentos) y no considera los otros tipos de recursos digitales (fotos, videos, software)

Se ha hecho estudios que demuestran que es una ventaja usar sistemas de recomendación orientados a la pedagogía y que usar ontologías genera más satisfacción que sólo usar sistemas basados en *ratings* (Nadolski et al., 2009). Hay que considerar que en el experimento realizado en el estudio no se usa data real y para simplificar se hacen

suposiciones e inicializaciones aleatorias en la simulación. El estudio también muestra las ventajas de utilizar un sistema híbrido en recomendación para educación, y que una recomendación basada en alguna ontología, llamado un conjunto limitado de características de los objetos y de los usuarios, debería ser parte del híbrido. Sin embargo, esto puede inducir gran complejidad al algoritmo, trabajo que se puede simplificar reduciendo el vocabulario de metadata, lo cual no es trivial pues no hay manera de saber exactamente qué conjunto de metadata limitada elegir. Es por esto que queremos comparar si realmente la metadata curricular y pedagógico-tecnológica son un aporte o si basta el contenido de los recursos para llegar a buenos resultados.

El uso de ontologías también se ha visto en otras investigaciones (Becerra et al., 2012) donde utilizando un modelo del dominio (ontología) y métodos de categorización se logra reducir el esfuerzo requerido para encontrar, reutilizar y crear objetos digitales de aprendizaje. Para mejorar el proceso de recomendación utilizan algoritmos de categorización basados en redes bayesianas, logrando, a través de ellos, enriquecer la descripción de los objetos digitales.

También podemos encontrar sistemas basados en memoria y basados en modelos, como los dos mencionados a continuación.

(Drachsler, Hummel, & Koper, 2008a) proponen una combinación de técnicas de recomendación basadas en memoria que parecen adecuadas para realizar recomendaciones personalizadas en actividades de aprendizaje en el contexto de *e-learning*, dando un modelo inicial del diseño de dicho sistema en redes de aprendizaje. La principal ventaja de estos sistemas de recomendación es que usan información extraída de los *ratings*, son independientes del dominio y no requieren análisis de contenido. Pero tienen a su vez una gran cantidad de desventajas, una de ellas es el llamado problema "*Cold-start*" ya que las técnicas colaborativas dependen de tener recolección de comportamiento del usuario en el pasado, es por eso que esto se produce al inicio del sistema de recomendación o cuando

un nuevo usuario o ítem entra al sistema. Otro problema es la escalabilidad ya que se trata con muchos volúmenes de datos que deben ser procesados para cada recomendación.

(Vialardi, Bravo, Shafti, & Ortigosa, 2009) proponen un sistema de recomendación basado en minería de datos para ayudar a los estudiantes a escoger los siguientes cursos a tomar en una Universidad. El sistema obtiene reglas que luego usa para decirle al usuario si el inscribirse en un curso tiene buena probabilidades de éxito o no. Esto ayuda al proceso de aprendizaje haciendo más fácil la decisión de los estudiantes sobre el camino curricular que seguir. Otros varios ejemplos donde se usan técnicas de minería de datos como clasificación, *clustering*, asociación y patrones en educación se pueden ver en (Adomavicius, Sankaranarayanan, Sen, & Tuzhilin, 2005).

Por otra parte, como se defina la información de los distintos recursos influye en el sistema de recomendación. (Vuorikari & Koper, 2009a) muestra que las búsquedas de los usuarios son más eficientes con estrategias de *Social Information Retrieval*. Traen mejores resultados con menos esfuerzo por parte del usuario. Además muestra que la navegación por internet en una comunidad no es suficiente para ayudar a los usuarios a descubrir una gran variedad de recursos fuera de la comunidad. Eso sí, utilizando *tags* se puede crear conexiones entre repositorios permitiendo una mayor reutilización entre ellos, incluso entre recursos de distintos idiomas. (Dron et al., 2000) también consideran folksonomías (*tags*) para realizar las recomendaciones, presentan el prototipo CoFind e indican que las folksonomías son importantes para generar categorías, más allá de las preferencias del usuario.

Uno de los grandes problemas de los sistemas de recomendación en educación es que no pueden compararse fácilmente unos con otros pues no hay muchos repositorios dedicados (Verbert et al., 2011), no así como en otros temas (películas, libros, música, etc.). Por esto es muy importante que el sistema de recomendación implementado se pruebe con data real y que la parametrización sea adecuada al escenario donde se quiere

implementar. (Manouselis et al., 2007) es un ejemplo de esto, el sistema de recomendación multi-atributos que proponen fue probado en un repositorio real y parametrizado para el caso específico que se estaba estudiando. También para esto (Drachsler, Hummel, & Koper, 2008b) propone la simulación de un modelo que permite evaluar los sistemas de recomendación simulando lo que normalmente ocurre cuando un estudiante desea lograr una meta en una red de aprendizaje informal.

2.3. Tipos de Información para Describir Recursos Digitales de Aprendizaje

Como mencionamos en la sección anterior es muy importante la forma de describir los recursos digitales para que los sistemas de recomendación puedan mejorar el descubrimiento y utilización de los recursos por parte de los profesores.

Un recurso digital de aprendizaje puede describirse de varias formas. La más fácil y muchas veces la más usada, es el contenido propio del recurso, por ejemplo el texto que compone un archivo, la descripción de una fotografía, el resumen de un sitio web, etc. Además, de esta forma, un recurso digital de aprendizaje se puede describir a través de metadata curricular identificando principalmente a qué nivel educacional corresponde, qué unidad está evaluando, tipo de enseñanza, etc. Por otra parte, la metadata que describe a un recurso digital de aprendizaje puede ser de tipo pedagógico-tecnológica (Mishra & Koehler, 2006), es decir, centrada en indicar cuál es el objetivo pedagógico del recurso, cuál es la audiencia objetivo, cuáles son sus características técnicas, si es un software, si es una fotografía jpg o png, si es una animación, etc.

Dentro de la metadata el estándar más conocido es LOM (Technology & Committee, 2002). Este estándar especifica la sintaxis y semántica de la metadata de los objetos de aprendizaje, los cuales son definidos como entidades, digitales o no digitales, que pueden ser usadas, reutilizadas o referenciadas durante aprendizaje apoyado por tecnología. El estándar de metadata de los objetos de aprendizaje se centra en el mínimo conjunto

de atributos necesarios para permitir que estos objetos sean manejados, ubicados y evaluados. Algunos atributos relevantes incluyen: tipo del objeto, autor, dueño, términos de distribución, formato. Puede también incluir atributos pedagógicos como estilo de enseñanza, nivel, pre-requisitos, etc.

A pesar de ser una forma estándar de describir recursos digitales de aprendizaje utilizar LOM tiene sus desventajas (Di Nitto et al., 2006): El significado exacto de cierta metadata no está claramente especificado, alguna metadata no es completamente procesable por máquinas e incluso algunas características de los recursos no pueden ser expresadas (no se puede decir si un recurso está enfocado a grupos de estudio o para estudio individual). Muchos estudios encuentran que la metadata existente es insuficiente, incluso un estudio reciente muestra que metadata LOM es típicamente incompleta, inexacta y no interpretable por una máquina (Friesen, 2004). De todas formas se ha propuesto extensiones de LOM (Di Nitto et al., 2006) para proveer aún más apoyo a otras actividades, como por ejemplo: configuración automática de software requerido, configuración automática de software soportado, tutoría (tutoría personalizada automática para estudiantes, recomendándole según sus características el mejor recurso), evaluación. Pero estas mejoras no resuelven todas las críticas.

Una ventaja de la metadata es que se puede incluso generar automáticamente. (Riley et al., 2009) describen la implementación de un *framework* para la generación automática de metadata de uso empírico. Lo que hacen es usar los *logs* para generar metadata sobre el uso de los recursos digitales de aprendizaje, agregando así más a la metadata estática de los recursos y de los usuarios, utilizando para esto selección de características (*feature selection*). La metadata permite desacoplarse del sistema de administración de los recursos digitales y de los recursos mismos, ya que permite que el contenido sea interoperable, mantenible y (en teoría) descubrible y reusable. De todas formas no se

hace una comparación con los otros tipos de información para indicar que efectivamente la metadata es la mejor forma para describir un recurso digital de aprendizaje.

Otra forma de describir los recursos digitales es a través de *tags* (Vuorikari & Põldoja, 2009). Hay muchos sistemas de *tags* dependiendo de la motivación de los usuarios para incluirlos, ya sea de uso personal, para compartir con la comunidad, etc. En (Vuorikari & Põldoja, 2009) se compara tres sistemas de *tags* encontrando diferencias cuando son de uso personal, donde se busca encontrar más de lo mismo o donde los *tags* son parte de la definición del contenido y la motivación es compartir el recurso con otros usuarios y llamar la atención en la nube de *tags*. A pesar de estas diferencias hay similitudes que permiten su reutilización en varios sistemas. Una de las similitudes notable es que cubren un número de tópicos que se comparten en varios sistemas educaciones (matemáticas, ciencias, etc.) dado que las audiencias objetivo son las mismas. Además hay *tags* que se pueden reutilizar en repositorios de distintos países dado que se escriben de manera similar en distintos idiomas, como nombres propios, países, acrónimos.

Los *tags* traen beneficios como flexibilidad, rápida adopción y fácil usabilidad (Krestel & Fankhauser, 2012), pero tienen algunas desventajas, sus significados tienen que ser evaluados e inferidos en el contexto de otros *tags* e información del usuario, *tags* incluso pueden tener significado sólo para el usuario que los creó, además pueden tener varios propósitos (referidos a anotar el recurso, referidos al usuario). El que usuarios puedan anotar libremente *tags* hacen que el sistema se llene de ruido y sea disperso.

A pesar de la existencia de estas formas de descripción de los recursos digitales de aprendizajes no hay estudios comparativos específicos que indiquen cuál de ellas es mejor para permitir el buen descubrimiento y utilización de los recursos. Tampoco se hace comparación entre los distintos tipos de metadata (curricular y pedagógica-tecnológica), por lo que vemos importante encontrar, para el escenario específico que estamos analizando, cuál de ellas es la más apropiada.

3. METODOLOGIA

Es muy difícil encontrar *data sets* que se hayan hecho públicos en *e-learning*. La mayoría de los sistemas de recomendación son probados en repositorios sintéticos o artificiales (creados para probar los sistemas, no representantes de un escenario real) o en repositorios representativos de otros dominios como películas, libros, productos a la venta, etc. Esto lleva a que luego, probando los mismos sistemas en escenarios reales de educación, se obtengan resultados completamente distintos (Verbert et al., 2011).

El escenario de *e-learning* presenta sus propias características (Drachsler et al., 2010), influyendo si el escenario es formal o informal, las características del *data set*, los modelos de los usuarios, las tareas y algoritmos de recomendación. (Sicilia et al., 2010) es un ejemplo de cómo algoritmos con buenos resultados en *data sets* normales tuvieron malos resultados en un repositorio de enseñanza aprendizaje (MERLOT), a pesar que hayan sido recomendaciones de calidad.

Es por lo anterior que buscamos un *data set real*, con recursos digitales de aprendizaje, donde profesores hicieran búsquedas para descubrir nuevo material que utilizar al preparar sus lecciones. Además, necesitábamos un sistema que no fuera intrusivo a la hora de recolectar las preferencias de los usuarios, ya que queríamos encontrar patrones que se dieran en la búsqueda y descubrimiento de recursos y por esto nos interesaba evitar el efecto "users-do-not-like-to-rate" (a los usuarios no les gusta calificar los recursos) (Nadolski et al., 2009).

Así, conociendo las características del repositorio a utilizar, en primer lugar hicimos un análisis teórico sobre qué tipo de sistema de recomendación se adecúa mejor a este escenario educativo. Luego, hicimos un análisis experimental preliminar, evaluando estos tipos de sistemas de recomendación para obtener el que mejor resultados presentara en cuanto a precisión y *recall*. Finalmente, para ver la influencia del tipo de información

utilizada para describir los recursos digitales, creamos siete bases de datos distintas, donde el único cambio entre ellas era el tipo de información utilizada para describir los recursos: contenido, metadata curricular, metadata pedagógico-tecnológica y todas sus combinaciones. Luego, evalúamos el mejor sistema de recomendación encontrado en cada una de las bases datos, pudiendo encontrar con qué tipo de información como descripción de los recursos se obtenían los mejores resultados.

Para realizar los experimentos utilizamos como conjunto de datos *logs* de visitas a recursos digitales de aprendizaje recolectados del portal EducarChile desde el año 2007 al año 2011. Estos *logs* incluían todos los accesos de todos los usuarios que alguna vez habían accedido al portal, usuarios registrados y no registrados, profesores, estudiantes, familia, etc. Por esto se efectúo primero una limpieza de los datos. Esta limpieza consistió en dejar sólo los *logs* correspondientes a profesores o estudiantes avanzados de pedagogía que habían visitado recursos digitales de aprendizaje. Además, se redujeron los datos a aquellos usuarios que habían visitados más de 10 recursos digitales de aprendizaje, para tener la confianza de que eran usuarios recurrentes. Luego de aplicar estos filtros nos quedamos con una base de datos de 25544 *logs*, 1376 usuarios y 43092 recursos digitales de aprendizaje.

Esta base de datos contenía, por cada recurso digital de aprendizaje: el contenido que los describe, en caso de recursos como imágenes, sitios web, etc. se cuenta con una descripción de su contenido; metadatos curriculares como la asignatura y nivel al que están enfocados; y otros metadatos pedagógico-tecnológicos como objetivo de aprendizaje, audiencia objetivo, formato del recurso. Para limpiar el contenido se eliminaron palabras comunes como artículos, conjunciones, etc. y se hizo un pre-procesamiento del texto, donde se ejecutaron algoritmos de *stemming* y lematización, con el fin de disminuir las variaciones de palabras y reducirlas a una raíz común (correr, corrieron y corren se reduce

a correr). Para ello se utilizó la librería *freeling* por sus buenos resultados al procesar texto en español.

Para la evaluación de todos los experimentos utilizamos validación cruzada 10 veces, es decir, dividimos los datos en 10 conjuntos distintos y consideramos, iterativamente, un conjunto como el conjunto de pruebas y el resto como el conjunto de entrenamiento.

4. ANALISIS PRELIMINAR

4.1. Sistema de Recomendación Adecuado para el Escenario Educativo del Portal EducarChile

En esta sección haremos una comparación de distintos tipos de sistemas de recomendación, como los que hemos descrito en las secciones anteriores, para encontrar el que mejor se adecúe al escenario educativo del repositorio EducarChile. Como vimos este sistema de recomendación debe considerar las siguientes tareas (Manouselis et al., 2011)

(i) Encontrar recursos interesantes y de buena calidad.

(ii) Encontrar recursos novedosos: esto significa explorar técnicas de recomendación que seleccionen ítems más allá de sus similitudes.

(iii) Encontrar compañeros: explorar atributos para medir la similitud con otras personas.

(iv) Encontrar buenos caminos de aprendizaje: explorar criterios para la construcción y sugerir secuencias alternativas (pero similares).

Ya que una de las tareas es encontrar recursos novedosos, no podemos utilizar un sistema de recomendación puramente basado en contenido, puesto que solamente entregará recomendaciones similares a recursos de aprendizaje previamente considerados por el educador. Dado que queremos ver cómo debe ser la descripción de los recursos de aprendizaje necesitamos considerar un filtrado basado en contenido pero incluyendo las ventajas del filtrado colaborativo, para superar los problemas de sobre-especialización y el ingreso de usuarios y recursos nuevos al sistema.

Por lo anterior, podemos decir que un sistema de recomendación híbrido es una buena elección. Pero como vimos hay muchos tipos de híbridos y necesitamos encontrar el que mejor se adecúe a las características de nuestro escenario educativo.

Por otro lado, nos interesa explicar cómo está funcionando el sistema de recomendación, qué patrones está encontrando, grupos de usuarios, de recursos, etc. Es por esto, que un sistema basado en modelos probabilísticos parece una buena opción. Con probabilidades se podrían encontrar varias relaciones dentro de los recursos digitales de aprendizaje. Dentro de estos sistemas, los modelos bayesianos se han usado desde hace ya varios años, siendo algunos ejemplos (Chen & George, 1999; Condliff et al., 1999). En (Breese, Heckerman, & Kadie, 1998) los resultados muestran que para un amplio rango de condiciones, las Redes Bayesianas con árboles de Decisión en cada nodo y métodos de Correlación superan métodos como *Baysian-clustering* y de Similitud de Vectores. Entre correlación y Redes Bayesianas, el método preferido depende de la naturaleza del conjunto de datos, naturaleza de la aplicación (*ranking*, presentación uno a uno), y la disponibilidad de evaluaciones de los usuarios con las cuales hacer predicciones. Cuando hay relativamente pocas evaluaciones, Correlación y Redes Bayesianas tienen menos ventaja sobre las otras técnicas.

Junto con Redes Bayesianas se han utilizados métodos de aprendizaje de máquina como *Singular Value Decomposition* (SVD) (Billsus & Pazzani, 1998) donde usan SVD de una matriz inicial de *ratings* de usuarios, explotando la estructura latente que esencialmente elimina la necesidad para usuarios de calificar ítems comunes para llegar a ser predictores de las preferencias de los otros usuarios. Encontrando que son mejor que los basados en Correlación.

Otro enfoque es *Probabilistic Matrix Factorization* (PMF) que tiene ventajas sobre todos los anteriores. Este método escala linealmente con el número de observaciones y tiene un buen desempeño en conjunto de datos grandes, escasos y no balanceados (Salakhutdinov & Mnih, 2007). Además en (Salakhutdinov & Mnih, 2008) presentan un tratamiento bayesiano al modelo PMF, donde la capacidad del modelo es controlada automáticamente integrando sobre todos los parámetros e hipermarámetros del modelo.

Muestran que los modelos bayesianos PMF, BPMF (Figura 4.1.a) pueden ser entrenados eficientemente usando métodos de Cadenas de Markov Monte Carlo, obteniendo mayor precisión en la predicción que modelos PMF entrenados usando estimaciones MAP.

Los sistemas BPMF funcionan bien para recomendar recursos conocidos pero no pueden generalizar para recursos no conocidos (Salakhutdinov & Mnih, 2008). En esto último es mejor LDA (D. M. Blei et al., 2003). LDA (Figura 4.1.b) es un modelo probabilístico generativo para colecciones de data discreta como texto. Es un modelo bayesiano jerárquico de tres niveles, en el cual cada ítem de la colección es modelado como una mixtura finita sobre un conjunto de tópicos subyacente. Cada tópico es, a su vez, modelado como un mixtura infinita sobre un conjunto de probabilidades de tópicos subyacente. En el contexto de modelamiento de texto, las probabilidades de tópicos dan una representación explícita de un documento.

Utilizar un modelo probabilístico como LDA, en general cualquier modelo que requiera una fase de aprendizaje, trae consigo algunos problemas de desempeño para entrenar el modelo. Por ejemplo, métodos de muestreo de Gibbs distribuído para LDA llegan a cuellos de botella por memoria o comunicación. Pero hay desarrollos que buscan mejorar y resolver estos problemas, como (Y. Wang, Bai, Stanton, Chen, & Chang, 2009) que propone estrategias para mejorar la escalabilidad, demostrando con sus experimentos que se reduce el cuello de botella de la comunicación no paralelizable y mejora la escalabilidad del LDA.

Una extensión del LDA es el *User Rating Profile* (URP) (Marlin, 2003), que introduce los *ratings* como parte del modelo para tener una distribución de *ratings* y no ser usado sólo como consumo (el usuario ha visto o no un recurso, donde no importa la evaluación o *rating* que le ha asignado) como en el caso del LDA. Incluso se han hecho mejoras utilizando muestreo de Gibbs (Barbieri, 2011) para la estimación de parámetros mostrando mejora en

la precisión además del tiempo de aprendizaje con respecto a la versión variacional. Como en nuestro caso utilizaremos visitas de los usuarios, no es necesario un modelo como URP.

LDA se ha aplicado en varios escenarios. (Fields, Rhodes, & Inverno, 2010) presenta un medio de comparación de listas de reproducción en un espacio de dimensión reducida a través del uso de nube de *tags* agregados y modelos temáticos. Utilizan LDA ocupando las nubes de *tags* de canciones como documentos y *tags* como las palabras. Así, el sistema toma los *tags* de todas las canciones, crea un modelo LDA que describa la distribución de los tópicos, infiere mixturas de tópicos para todas las canciones y crea base de datos de vectores de listas de reproducción. En (Krestel & Fankhauser, 2012) introducen un sistema de recomendación de *tag*s que combina modelos probabilísticos de *tags* provenientes del recurso con *tags* provenientes del usuario utilizando LDA, lo que les da la posibilidad de recomendar *tags* no previamente asignados al recurso o usados por el usuario. Los resultados muestran que LDA ayuda, en particular, para generalizar prácticas individuales de *tagging* en recursos.

Con los antecedentes anteriores vemos que un modelo probabilístico como LDA sería apropiado para formar parte del sistema híbrido de recomendación. En la literatura encontramos un sistema de recomendación híbrido que ocupa LDA, además de métricas de filtrado colaborativo, llamado CTR (C. Wang & Blei, 2011). Este sistema (Figura 4.1.c) es utilizado para recomendar artículos científicos, permitiendo realizar recomendaciones de artículos existentes y recientemente publicados. Además, da representaciones interpretables tanto de usuarios como de ítems. Para realizar esto, combina técnicas de filtrado colaborativo basados en modelos de factores latentes (Salakhutdinov & Mnih, 2008) y análisis de contenido basado en modelos probabilísticos de tópicos (D. M. Blei et al., 2003). Estas técnicas se combinan en un modelo probabilístico donde hacer recomendaciones para un usuario en particular es parecido a calcular una esperanza

condicional de variables ocultas. El algoritmo para calcular estas esperanzas balancea naturalmente la influencia del contenido de los artículos y preferencias de otros usuarios.

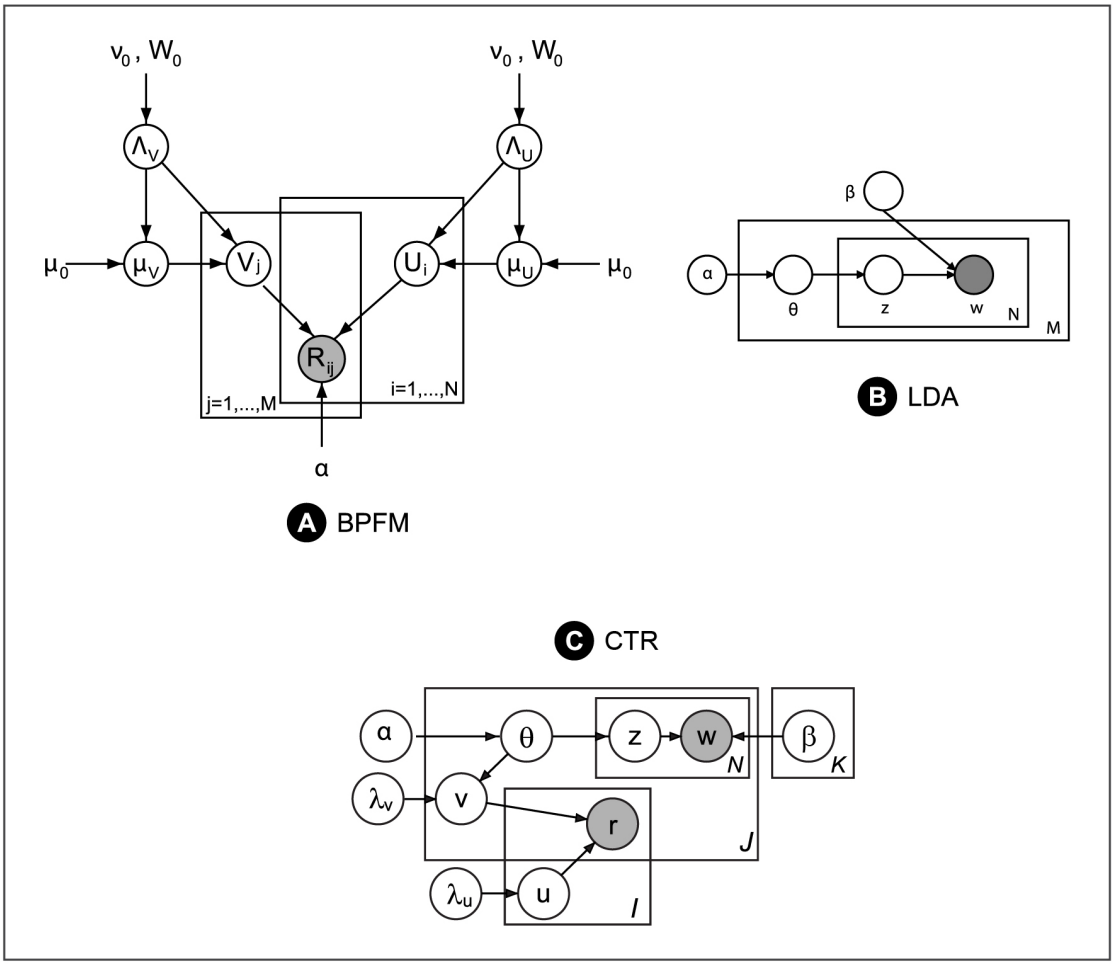

FIGURA 4.1. Modelo Gráfico de Sistemas de Recomendación Probabilísticos Comparados en esta Investigación

En la Figura 4.1 se muestran los tres sistemas de recomendación a comparar, luego del análisis realizado previamente.

En BPMF aplicado al contexto de esta investigación N sería el número de usuarios y M el número de recursos digitales. Así R_{ij} representa el valor del *rating* dado por el usuario i al recurso j. U_i y V_j representan vectores D-dimensionales de características latentes (*latent features*) específicos para el usuario y para el recurso respectivamente.

En LDA aplicado al contexto de esta investigación N es el número de palabras que representa el recurso digital de aprendizaje, M el número de recursos y K es el número de tópicos. En cuanto a las variables, w representa las palabras que representan el recurso, z representa los tópicos latentes, β es una matriz donde $\beta_{ij} = p(w_j = 1/z_i = 1)$, α es parámetro a nivel de grupo de recursos y $\theta \sim Dir(\alpha)$.

En CTR aplicado al contexto de esta investigación se mantiene la interpretación del LDA en la parte que reutiliza. Además, u representa el vector de usuarios donde $u_i \sim N(0, \lambda_u^{-1} I_k)$, v representa el vector latente de un recurso donde $v_j = \epsilon_j + \theta_j$ y $\epsilon_j \sim N(0, \lambda_v^{-1} I_k)$, r representa el *rating* y $r_{ij} \sim N(u_i^T v_j, c_{ij}^{-1})$.

En la siguiente sección se evaluará estos sistemas de recomendación para encontrar el que mejor se adecúe al conjunto de datos de la investigación

5. RESULTADOS

5.1. Evaluación Sistema de Recomendación Educacional Colaborativo, por Contenido e Híbrido

Ya que la operación exitosa de los sistemas de recomendación en el contexto de aplicaciones de la vida real requiere de un cuidadoso testing y parametrización antes de su puesta en marcha (Manouselis et al., 2010), comparamos el sistema híbrido CTR, que teóricamente debería ser el más adecuado, con un sistema de recomendación colaborativo clásico (utilizando Correlación de Pearson), con PMF, BPMF y con LDA.

En primer lugar comparamos los sistemas de recomendación que utilizan estimación de *ratings* para hacer las predicciones, que en este caso serían (1) Filtrado colaborativo con Correlación de Pearson, (2) PMF y (3) BPFM. Para PMF y BPMF el modelo fue aprendido con el conjunto de entrenamiento y sus resultados evaluados en el conjunto de test.

En la Tabla 5.1 se muestra los resultados de desempeño para cada uno de ellos. Las medidas de evaluación fueron calculadas tal como se describen en (Herlocker et al., 2004).

De los resultados podemos ver que el sistema de recomendación con menor error y mayor precisión es BPMF, teniendo un error NMAE de 0.153 y una precisión de 0.493. Siendo mejor que PMF y que filtrado colaborativo con correlación de Pearson para este escenario particular. De los restantes, PMF obtiene una mejor precisión que Person, con un 0.211, sin embargo, Pearson obtiene un menor error NMAE que PMF con 0.164.

Luego comparamos BPMF con los demás sistemas que no utilizan estimación de *ratings* para hacer las recomendaciones, LDA y CTR. Para LDA el parámetro α se estimó junto con el aprendizaje del modelo, y los tópicos se inicializaron a partir de una distribución suavizada desde recursos elegidos aleatoriamente. Para CTR se utilizó $a = 1$, $b = \lambda_u = 0.01$ y $\lambda_v = 10$ (llegando al valor de λ_v luego de experimentar con valores de $\{10, 100, 1000$ y $10000\}$ y viendo en cual escenario se obtenían mejores resultados).

TABLA 5.1. Comparación Sistemas de Recomendación que Utilizan Estimación de Ratings

	MAE	**RMSE**	**NMAE**	**Precision**
BPMF	0.612	0.829	0.153	0.493
PMF	0.739	1.040	0.185	0.211
Pearson	0.818	1.233	0.164	0.194

Además, el número de tópicos (K) también fue determinado luego de varios experimentos para encontrar qué valor permitía los mejores resultados, llegando a K = 100.

Como no hay estimación de *ratings*, y además como las evaluaciones en cero (no visitada) son inciertas (pueden indicar que un usuario no le gusta un artículo o que no sabe de su existencia), utilizamos *recall-M* (C. Wang & Blei, 2011).

$$recall - M = \frac{número\ de\ artículos\ que\ le\ gustan\ al\ usuario\ en\ top\ M}{total\ de\ artículos\ que\ le\ gustan\ al\ usuario}$$

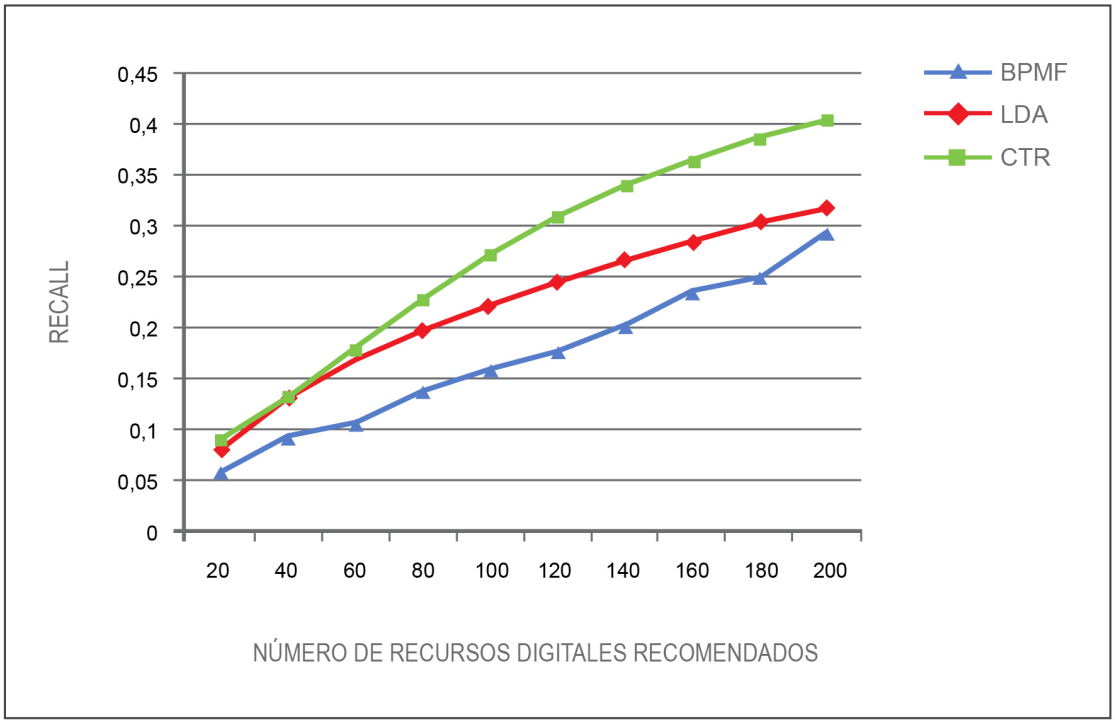

FIGURA 5.1. Comparación Sistemas de Recomendación Probabilísticos

Un sistema con alto *recall-M* y un M menor será un mejor sistema. Esto nos permite medir *recall* y precisión (mientras más pequeño M más preciso es el sistema de recomendación).

En la Figura 5.1 podemos notar que CTR es el mejor de los tres sistemas de recomendación, en cuanto a *recall-M*, para el escenario de esta investigación. A partir de 60 recursos digitales recomendados se puede notar una mayor diferencia entre los tres sistemas de recomendación, aumentando esta diferencia a medida que aumenta el número de recursos recomendados.

Con lo anterior podemos ver que un sistema de recomendación híbrido, que considera el contenido de recursos digitales y las preferencias de usuarios similares, es el que mejor precisión tiene en las recomendaciones de un repositorio de educación como EducarChile, además de tener la capacidad de abarcar más recursos para recomendar (*recall*), mejorando así la recomendación de recursos nuevos. Además, podemos ver que utilizar un algoritmo basado en modelos (como lo es CTR) es favorable en este escenario, ya que se puede aprender un modelo a partir de los datos, que permite luego realizar mejores recomendaciones.

5.2. Encontrando la Metadata en nuestro Conjunto de Datos Educacional

Para encontrar el tipo de información que mejor describa a los recursos digitales de aprendizaje del portal educacional EducarChile se crearon siete conjuntos de datos a partir del conjunto de datos inicial, variando solamente el tipo de información que describía a los recursos digitales:

(i) Información del recurso digital es el contenido mismo del recurso (texto o descripción en caso de imágenes, sitios educativos, software, etc.).

(ii) Información del recurso digital corresponde a los metadatos curriculares que lo describen.

(iii) Información del recurso digital corresponde a los metadatos pedagógico tecnológicos que lo describen.

(iv) Información del recurso digital corresponde a su contenido además de los metadatos curriculares que lo describen.

(v) Información del recurso digital corresponde a su contenido además de los metadatos pedagógico tecnológicos que lo describen.

(vi) Información del recurso digital corresponde a los metadatos curriculares más los metadatos pedagógico tecnológicos que lo describen.

(vii) Información del recurso digital corresponde a su contenido además de los metadatos curriculares y pedagógico tecnológicos que lo describen.

Para evaluar en cuál de estos siete escenarios el sistema de recomendación CTR se comporta mejor utilizamos como métrica nuevamente el *recall-M*.

En la Figura 5.2 podemos ver que el sistema de recomendación mostró un mejor desempeño cuando los recursos digitales de aprendizaje eran descritos por metadata curricular, sobrepasando el 0.5 de *recall* a partir de los 80 recursos digitales recomendados, llegando hasta un 0.55 para 200 recursos, mientras que en ningún otro caso se llegó hasta un nivel tan alto. Por ejemplo, utilizando metadata curricular y pedagógico-tecnológica para describir a los recursos digitales de aprensizaje, el cual mostró el segundo mejor desempeño, sólo se llegó a un *recall* de 0.27 con 80 recursos llegando a un *recall* máximo de 0.41.

Así, al considerar únicamente el texto correspondiente a los metadatos curriculares (ej. matemáticas, lenguaje, 8 básico, etc.) y las visitas de los usuarios a recursos digitales (sistemas de recomendación híbridos), tanto la precisión en recomendar objetos, como el alcance (*recall*) de los mismos (ej. para recomendar nuevos objetos) mejora significativamente, es decir, los metadatos curriculares son buenos discriminadores y predictores del interés de los profesores en recursos educacionales.

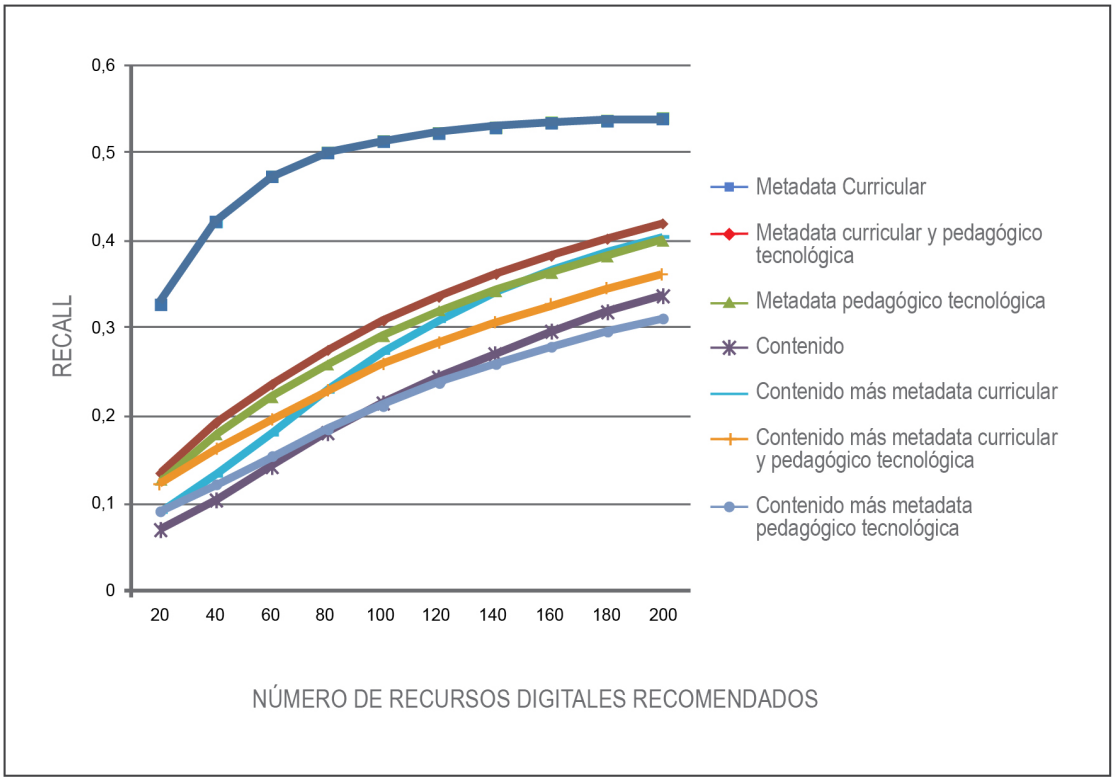

FIGURA 5.2. Comparación Tipos de Información para Describir Recursos Digitales de Aprendizaje

Podemos ver también que agregar metadata pedagógico-tecnológica no aporta de buena manera, sino que empeora el desempeño del sistema de recomendación, aunque por sí sola es mejor que utilizar sólo el contenido de los recursos como su descripción. Esto sucede porque la información contenida en los metadatos pedagógico-tecnológicos es redundante con la información contenida en los metadatos curriculares y/o en el contenido propio de los recursos. Los algoritmos probabilísticos utilizados penalizan las palabras que están repetidas en todos los recursos, ya que al aparecer muy repetidas en recursos distintos indican que son muy "populares" y por ende, no deben considerarse para indicar que dos recursos se parecen. Así, al incluir los metadatos pedagógico-tecnológicos, muchas palabras se estarían volviendo redundantes disminuyendo el desempeño del sistema de recomendación.

Por otra parte, si analizamos los grupos de recursos digitales que surgieron a partir del modelo aprendido de los datos podemos ver que hay grupos de recursos que no son cubiertos por los metadatos curriculares (Tabla 5.2), como por ejemplo “Artículos de actualidad sobre educación” o “Estrategias de enseñanza” o “Educación infantil”, entre otros. Aunque no ayudan directamente en preparar las lecciones, sí son recursos novedosos que podrían ayudar a los profesores a mejorar, por lo que debería considerarse metadata que describa estos tipos de recursos, ya que la metadata curricular los deja de lado.

TABLA 5.2. Ejemplos de Grupos de Información Generados no Considerados por Metadata Curricular

	Artículos más representativos del grupo
tópico 1	- Ventajas y desventajas del ítem de selección múltiple - Recomendaciones generales para la corrección de preguntas abiertas - Sugerencias de construcción de ítemes - La estrategia didáctica genérica - Revolucionario método de enseãnza para trastornos del lenguaje. La Nación - Definen las 10 claves del aprendizaje efectivo. El Mercurio - Sugerencias para el trabajo colectivo - Trabajo en grupo y trabajo individual - Profesores comienzan a probar estrategias de enseãnza en 140 caracteres. El Mercurio. - Principales métodos de una pedagogía activa
tópico 2	- El constructivismo en la Educación Infantil - La cooperación en Educación Infantil - Programas y líneas de acción de educación parvularia, 2006 - Reforma educacional chilena: descripción general y niveles de concreción curricular. - Más y mejor educación preescolar/ Harald Beyer. El Mercurio - Consideraciones sobre la educación para los niños trabajadores - La pedagogía Freinet en la Educación Infantil II - La importancia de invertir en la educación preescolar - De la cooperación en Educación Infantil y algunos de sus dilemas más controvertidos - Evaluación del Proyecto de Jardines Experimentales de Fundación Integra
tópico 3	- ¿Cómo está cambiando la educación con la puesta en marcha de la ley SEP? - Autoridades de Unesco y expertos en educación analizaron impacto de la crisis económica en sistemas educativos de Latinoamérica. La Tercera - Subsecretaria Romaguera: "tecnología en liceos y escuelas ha tenido positivos efectos como factor de equidad". Mineduc - Más de 400 mil niñas y niños más vulnerables tendrán mejor educación gracias a nueva ley sobre Subvención Preferencial. Mineduc - UNESCO presenta sistema para que las escuelas mejoren la atención a estudiantes con necesidades educativas especiales. UNESCO - UNESCO señala que Latinoamérica va adelantada hacia metas educativas establecidas para el 2015. La Tercera - Educación inicial de calidad asegura un correcto desarrollo durante la vida adulta. La Tercera - Expertos señalan cómo mejorar notas en enseñanza media. La Tercera - Concurso incentiva a profesores a crear y compartir sus planificaciones en internet. La Tercera - Experto en educación trae a Chile las claves para una buena escuela. La Tercera

6. CONCLUSIONES Y TRABAJO FUTURO

6.1. Revisión de los Resultados y Comentarios Generales

En esta investigación revisamos las características que deben tener los sistemas de recomendación para ayudar en el proceso de búsqueda y descubrimiento de recursos digitales de aprendizaje, además de cómo deben estar descritos estos recursos para ayudar al sistema de recomendación.

Evaluando distintos sistemas de recomendación, llegamos a la conclusión que, para el escenario del portal educacional EducarChile, un sistema de recomendación híbrido (CTR), que utiliza modelos temáticos como uno de sus componentes, es el que tiene un mejor desempeño (precisión y *recall*), en comparación a sistemas de recomendación colaborativos y por contenido.

Podemos concluir también que utilizar metadata curricular para describir los recursos digitales de aprendizaje, en este escenario educacional específico, ayuda a mejorar el desempeño de los sistemas de recomendación, y por ende a facilitar la búsqueda y descubrimiento de nuevos recursos. De todas formas pudimos ver, que había categorías de recursos que quedaban fuera si sólo se consideraba estos metadatos, por lo que hay que considerar las características propias de cada comunidad para definir metadata apropiada, en este caso, metadata que incluyera artículos de actualidad y contenido de mejora para los profesores, por ejemplo.

Es interesante encontrar que los metadatos significativos no son parte de la metadata oficial (LOM) y que las recomendaciones sobre el texto libre tampoco son suficientes como para generar recomendaciones significativas. Por otro lado, estos resultados analizan el comportamiento de una comunidad que se ve influenciada por los mecanismos de búsqueda del sitio web correspondiente, donde las búsquedas son implementadas a través del buscador Google y alguna información (ej. Artículos de actualidad, Artículos sobre

educación infantil, etc.) están ocultos en el primer nivel de la interfaz de usuario del sistema.

Creemos que este estudio es un comienzo para ahondar más en el tema de mejorar el descubrimiento y utilización de recursos digitales por parte de profesores. A pesar de llegar a conclusiones sólo en el escenario específico del portal educacional EducarChile es fácil interpolar los resultados a escenarios educacionales parecidos, como por ejemplo Latinoamérica, donde el escenario educativo de la Educación Básica y Media está fuertemente regulado (currículum fijo).

6.2. Temas de Investigación Futura

Como trabajo futuro sería interesante incluir en las evaluaciones a usuarios reales del sistema, utilizando encuestas o sistemas más personales, para poder medir no sólo desempeño y *recall*, si no también calidad de los recursos encontrados y novedad de lo sugerido.

Además sería interesante comparar los resultados con otras bases de datos de portales educativos de otros escenarios a lo largo de Latinoamérica, para ver si se puede generalizar más lo encontrado en esta investigación.

Otra contribución como trabajo futuro sería anonimizar y publicar el set de datos generado para esta investigación, a partir de los logs recopilados de EducarChile, para permitir que otros sistemas de recomendación puedan probarse en el mismo escenario.

BIBLIOGRAFIA

Aciar, S., Zhang, D., Simoff, S., & Debenham, J. (2007). Informed Recommender : Basing Recommendations on Consumer Product Reviews. *Intelligent Systems, IEEE*(June), 39–47.

Adomavicius, G., Sankaranarayanan, R., Sen, S., & Tuzhilin, A. (2005). Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach University of Minnesota University of Connecticut. *ACM Transactions on Information Systems*, *23*(1), 103–145.

Adomavicius, G., & Tuzhilin, a. (2005, June). Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749.

Barbieri, N. (2011, July). Regularized Gibbs Sampling for User Profiling with Soft Constraints. *2011 International Conference on Advances in Social Networks Analysis and Mining*, 129–136.

Becerra, C., Astudillo, H., & Mendoza, M. (2012). Improving Learning Objects Recommendation Processes by Using Domain Description Models. *Conferencias LACLO*.

Billsus, D., & Pazzani, M. J. (1998). Learning Collaborative Information Filters.

Blei, D., Carin, L., & Dunson, D. (2012, November). Probabilistic Topic Models. *IEEE Signal Processing Magazine*, 77–84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bobadilla, J., Serradilla, F., & Hernando, A. (2009, May). Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems*, *22*, 261–265.

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 43–52.

Burke, R. (2007). Hybrid Web Recommender Systems. *The Adaptive Web*, *4321*, 377–408.

Chen, Y.-h., & George, E. I. (1999). A bayesian model for collaborative filtering. *Proceedings of the 7th International Workshop on Artificial Intelligence and Statistics*.

Cho, Y. H., Kim, J. K., & Kim, S. H. (2002, October). A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications*, *23*(3), 329–342.

Clements, K., & Pawlowski, J. (2012, February). User-oriented quality for OER: understanding teachers' views on re-use, quality, and trust. *Journal of Computer Assisted Learning*, *28*, 4–14.

Condliff, M. K., Madigan, D., Lewis, D. D., & Posse, C. (1999). Bayesian Mixed-Effects Models for Recommender Systems.

Di Nitto, E., Mainetti, L., Monga, M., Sbattella, L., & Tedesco, R. (2006). Supporting Interoperability and Reusability of Learning Objects : The Virtual Campus Approach. *Journal of Educational Technology & Society*, *9*(2), 33–50.

Drachsler, H., Bogers, T., Vuorikari, R., Verbert, K., Duval, E., Manouselis, N., et al. (2010, January). Issues and considerations regarding sharable data sets for recommender systems in technology enhanced learning. *Procedia Computer Science.*, *1*(2), 2849–2858.

Drachsler, H., Hummel, H., & Koper, R. (2008b). Using Simulations to Evaluate the Effects of Recommender Systems for Learners in Informal Learning Networks. *Proceedings of the 2nd Workshop on Social Information Retrieval for Technology Enhanced Learning, SIRTELâ08 at the 3rd European Conference on Technology Enhanced Learning (EC-TEL 2008)*.

Drachsler, H., Hummel, H. G. K., & Koper, R. (2008a). Personal recommender systems for learners in lifelong learning networks : the requirements , techniques and model. *Journal of Learning Technology*.

Dron, J., Siviter, P., Boyne, C., & Mitchell, R. (2000). CoFIND: steps towards a self-organising learning environment. *Proceedings of WebNet*, 146–151.

Farzan, R., & Brusilovsky, P. (2004). AnnotatEd : A Social Navigation and Annotation Service for Web-based Educational Resources. (Brusilovsky 2001).

Fields, B., Rhodes, C., & Inverno, M. (2010). Using Song Social Tags and Topic Models to Describe and Compare Playlists. *Workshop on Music Recommendation and Discovery, ACM, Barcelona (2010)*.

Friesen, N. (2004). The International Learning Object Metadata Survey. *The International Review of Research in Open and Distance Learning*.

Gabrilovich, E., & Markovitch, S. (2006). Overcoming the Brittleness Bottleneck using Wikipedia : Enhancing Text Categorization with Encyclopedic Knowledge.

Gabrilovich, E., & Markovitch, S. (2007). Computing Semantic Relatedness using Wikipedia-based Explicit Semantic Analysis. *International Joint Conference on Artificial Intelligence*.

Herlocker, J. L., Konstan, J. a., Terveen, L. G., & Riedl, J. T. (2004, January). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, *22*(1), 5–53.

Hummel, H. G. K., Berg, B. V. D., Berlanga, A. J., Drachsler, H., Janssen, J., Nadolski, R., et al. (2007). Combining social-based and information-based approaches for personalised recommendation on sequencing learning activities. *International Journal of Learning Technology*, *3*(2).

Krestel, R., & Fankhauser, P. (2012, January). Personalized topic-based tag recommendation. *Neurocomputing*, *76*, 61–70.

Lops, P., Gemmis, M. D., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook* (pp. 73–105). Boston, MA: Springer US.

Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender Systems in Technology Enhanced Learning. In *Recommender systems handbook* (pp. 387–415).

Manouselis, N., Vuorikari, R., & Van Assche, F. (2007). Simulated Analysis of MAUT Collaborative Filtering for Learning Object Recommendation. *Proceedings of the 1st Workshop on Social Information Retrieval for Technology-Enhanced Learning & Exchange*.

Manouselis, N., Vuorikari, R., & Van Assche, F. (2010, July). Collaborative recommendation of e-learning resources: an experimental investigation. *Journal of Computer Assisted Learning*, *26*(4), 227–242.

Marlin, B. (2003). Modeling User Rating Profiles For Collaborative Filtering.

Middleton, S. E., Shadbolt, N. R., & De Roure, D. C. (2004, January). Ontological user profiling in recommender systems. *ACM Transactions on Information Systems*, *22*(1), 54–88.

Mishra, P., & Koehler, M. J. (2006, June). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, *108*(6), 1017–1054.

Nadolski, R. J., Van Den Berg, B., Berlanga, A. J., Drachsler, H., Hummel, H. G. K., Koper, R., et al. (2009). Simulating light-weight Personalised Recommender Systems in Learning Networks : A case for Pedagogy-Oriented and Rating-based Hybrid Recommendation Strategies. *Journal of Artificial Societies and Social Simulation*, *12*.

Nikolopoulos, G., Kalou, A., Pierrakeas, C., & Kameas, A. (2012). Creating a LO Metadata Profile for Distance Learning: An Ontological Approach. *Communications in Computer and Information Science*, *343*, 37–48.

Ono, C., Kurokawa, M., Motomura, Y., & Asoh, H. (2007). A Context-Aware Movie Preference Model Using a Bayesian Network for Recommendation and Promotion. *User Modeling*, *4511*.

Pegler, C. (2012). Herzberg , hygiene and the motivation to reuse : Towards a three-factor theory to explain motivation to share and use OER. *Journal of Interactive Media in Education*.

Pennock, D. M., Horvitz, E., Lawrence, S., & Giles, C. L. (2000). Collaborative Filtering by Personality Diagnosis : A Hybrid Memory- and Model-Based Approach. *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, 473–480.

Petrides, L., Nguyen, L., Kargliani, A., & Jimes, C. (2008). Open Educational Resources : Inquiring into Author Reuse Behaviors. *Times of Convergence. Technologies Across Learning Contexts*, *5192*, 344–353.

Ramos, J. (2003). Using TF-IDF to Determine Word Relevance in Document Queries.

Riley, S. A., Miller, L. D., Soh, L. K., Samal, A., & Nugent, G. (2009). Intelligent Learning Object Guide (iLOG): A Framework for Automatic Empirically- Based Metadata Generation.

Rumelhart, D., Hinton, G., & Williams, R. (2002). Learning Representations by Back-Propagating Errors. In *Cognitive modeling.*

Salakhutdinov, R., & Mnih, A. (2007). Probabilistic Matrix Factorization. *Advances in Neural Information Processing Systems*.

Salakhutdinov, R., & Mnih, A. (2008). Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. *Proceedings of the 25th international conference on Machine learning - ICML '08*.

Santos, O. C., & Boticario, J. G. (2010, January). Modeling recommendations for the educational domain. *Procedia Computer Science*, *1*, 2793–2800.

Sarwar, B., Karypis, G., Konstan, J., Riedl, J., Ã, O. U. O. O. Y. O. O. O. O. O. O., ÒóûÃ°, E. O. . N. . N. O. O. Y. O. . N. O., et al. (2001). Item-Based Collaborative Filtering Recommendation. *Proceedings of the 10th international conference on World Wide Web*.

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative Filtering Recommender Systems. *The Adaptive Web*, 291–324.

Shepitsen, A., Gemmell, J., Mobasher, B., & Burke, R. (2008). Personalized recommendation in social tagging systems using hierarchical clustering. *Proceedings of the 2008 ACM conference on Recommender systems - RecSys '08*, 259.

Sicilia, M.-A., García-Barriocanal, E., Sánchez-Alonso, S., & Cechinel, C. (2010, January). Exploring user-based recommender results in large learning object repositories: the case of MERLOT. *Procedia Computer Science*, *1*, 2859–2864.

Tang, T., & McCalla, G. (2004a). Beyond Learnersâ Interest: Personalized Paper Recommendation Based on Their Pedagogical Features for an e-Learning System. *PRICAI 2004: Trends in Artificial Intelligence*, *3157*, 301–310.

Tang, T., & McCalla, G. (2004b). Utilizing Artificial Learners to Help Overcome the Cold-Start Problem in a Pedagogically-Oriented Paper Recommendation System. *Adaptive Hypermedia and Adaptive Web-Based Systems*, *3137*, 245–254.

Technology, I. L., & Committee, S. (2002). LOM: Standard for Learning Object Metadata. *IEEE 1484.12.1*.

Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., & Duval, E. (2011). Dataset-driven Research for Improving Recommender Systems for Learning. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*.

Vialardi, C., Bravo, J., Shafti, L., & Ortigosa, A. (2009). Recommendation in Higher Education Using Data Mining Techniques. *Educational Data Mining*.

Vuorikari, R., & Koper, R. (2009a). Ecology of social search for learning resources. *Campus-Wide Information Systems*, *26*(4), 272 – 286.

Vuorikari, R., & Koper, R. (2009b). Evidence of cross-boundary use and reuse of digital educational resources. *International Journal of Emerging Technologies in Learning*, *4*(4).

Vuorikari, R., & Põldoja, H. (2009). Comparison of educational tagging systems - any chances of interplay ?

Wang, C., & Blei, D. M. (2011). Collaborative topic modeling for recommending scientific articles. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11*.

Wang, Y., Bai, H., Stanton, M., Chen, W.-y., & Chang, E. Y. (2009). PLDA : Parallel Latent Dirichlet Allocation for Large-scale Applications. *Algorithmic Aspects in Information and Management*, *5564*, 301–314.

Wolpers, M., Memmel, M., & Giretti, A. (2009). Metadata in architecture education - first evaluation results of the MACE system. *Learning in the Synergy of Multiple Disciplines*, *5794*, 112–126.

Yu, K., Schwaighofer, A., Tresp, V., Ma, W.-Y., & Zhang, H. (2002). Collaborative Ensemble Learning : Combining Collaborative and Content-Based Information Filtering via Hierarchical Bayes. *Proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence*, 616–623.