EVALUACIÓN DEL USO DE ALGORITMOS COLABORATIVOS PARA ORIENTAR ACADÉMICAMENTE AL ALUMNADO EN BACHILLERATO

Memoria Investigadora del
Segundo Año de Doctorado que presenta

Emilio José Castellano Torres

Dirigida por el profesor

D. Luís Martínez López

Para la obtención del

DIPLOMA DE ESTUDIOS AVANZADOS

Departamento de Informática
Universidad de Jaén
Septiembre de 2007

AGRADECIMIENTOS

Agradecimientos a Luis Martínez y a Manolo Barranco por su apoyo y ayuda para conseguir en un tiempo récord algo que, en mi cabeza, empezó siendo otra cosa, y que me ha abierto un nuevo horizonte a explorar.

Quiero agradecer la desinteresada ayuda que me han ofrecido María José F. y Miguel Ángel T., porque aunque no lo sepáis hasta ahora, sin vosotros este trabajo no habría sido posible.

También a Juan, por haberse interesado y ofrecido en tantas ocasiones; y a los que han creído y creen en mí, aunque no me lo hayan dicho.

Me gustaría agradecer a Manuel su apoyo logístico y sobre todo moral, pues siempre ha sido de gran ayuda, y que en la mayoría de las veces no se ha ceñido al reducido ámbito de este trabajo, aunque también.

Y sobre todo, a mi padre y a mi madre, porque siempre han estado ahí para todo…

1. INTRODUCCIÓN

1 INTRODUCCIÓN

Las personas se enfrentan diariamente con situaciones en las que deben tomar decisiones, más o menos importantes, y se encuentran frente a una gran cantidad de opciones entre las que elegir, de modo que cada vez más se necesita una ayuda externa a la hora de explorar o filtrar dichas posibilidades, tanto para ganar tiempo como para mejorar la calidad de la decisión tomada.

Los Sistemas de Recomendación se han convertido en herramientas extremadamente útiles al proporcionar ayuda a la hora de tomar decisiones en una innumerable cantidad de ámbitos, con gran número de aplicaciones comerciales que abarcan ocio, compras, economía… [1-4]. Hoy día podemos entrar en Internet y casi sin darnos cuenta ser ayudados a elegir una o varias de entre innumerables posibilidades, sin importar el campo de actuación: tenemos por ejemplo buscadores que no sólo buscan páginas web en base a ciertos criterios, sino que internamente realizan una criba mostrándonos sólo las que creen de más interés para nosotros; y desde ahí podemos encontrar una vasta cantidad de recomendadores que nos proporcionan información útil a la hora de elegir qué producto comprar, qué restaurante visitar, qué música escuchar, qué libro leer, qué película ver, y un largo etcétera que se pierde en la profundidad de la Red.

Nadie está dispuesto a tomar una decisión a ciegas, por muy irrelevante que pueda parecer dicha decisión, y estos sistemas de recomendación son una excelente ayuda que suele mostrarse tan rápida como fiable, y que incluso sin perder de vista todas sus limitaciones se muestran como herramientas que marcan diferencia.

1.1 Sistemas de Recomendación

Por lo general, los sistemas de recomendación se encargan de proporcionar a los usuarios consejos e información personalizada sobre productos o servicios que puedan ser de interés a la hora de tomar una decisión. Este proceso, en el que el sistema guía al usuario a la hora de realizar una elección, puede proporcionar resultados que sean de gran utilidad, ya sea ahorrando tiempo, proporcionando datos relevantes de forma cómoda y fácil, e incluso obteniendo información que permite valorar opciones que de otra forma antes no se habrían contemplado, algo muy apreciado por la mayoría de los usuarios.

Simplificando, podemos decir que un sistema de recomendación reduce su problemática a predecir la puntuación que el usuario daría a una serie ítems que todavía no ha puntuado. Esta predicción, de la forma más intuitiva, podría basarse en las puntuaciones que otros usuarios asignaron en el pasado (sistemas colaborativos o sistemas de filtrado colaborativo), aunque también puede obtenerse usando otra clase de información distinta a las valoraciones

mencionadas, como pueden ser las características concretas de los productos, los perfiles de preferencias de los distintos usuarios a la hora de tomar decisiones, información demográfica sobre los usuarios, heurísticas basadas en comportamientos humanos, etc., y por supuesto, usar combinaciones de diversas aproximaciones de entre las existentes.

Lógicamente, en una puntuación o valoración lo que un usuario expresa es su grado de aceptación, su nivel de gusto o disgusto hacia un ítem concreto. De esta forma, y una vez estimadas las puntuaciones para aquellos ítems que el usuario no había valorado, podemos recomendar al usuario el mejor de entre ellos, es decir, el objeto o los objetos que mayor puntuación estimada obtuvieron.

Aunque todos los sistemas de recomendación tienen el mismo objetivo, ayudar al usuario realizando una serie de recomendaciones de forma que se simplifique al máximo la búsqueda que el usuario debe realizar, existen diversas fuentes de las que se nutren de la información requerida y diversas técnicas mediante las que estos sistemas calculan, evalúan, construyen y proporcionan sus resultados [3, 5, 6]:

- **Sistemas de recomendación colaborativos** [5, 7-13]: es el tipo más conocido de sistemas de recomendación. Al usuario se le recomendarán aquellos elementos elegidos anteriormente por gente con preferencias y gustos similares. Se puede decir que éstos son los primeros sistemas de recomendación que utilizan *estereotipos* como mecanismo para construir modelos de usuarios basándose en una cantidad limitada de información sobre cada usuario.
- **Sistemas de recomendación basados en contenido** [12, 14, 15]: al usuario se le recomendarán ítems parecidos a aquellos que eligió anteriormente; muchos sistemas basados en contenido se centran en la recomendación de ítems contenedores de información textual. Estas técnicas utilizan perfiles con información relativa a los usuarios, sus gustos, preferencias y necesidades.
- **Sistemas de recomendación demográficos** [12, 16]: clasifican a los usuarios en grupos demográficos basándose en ciertos atributos personales cuya información previamente se ha recolectado, y proporcionan recomendaciones potencialmente interesantes para cualquier persona perteneciente a dicho grupo demográfico.
- **Sistemas de recomendación basados en conocimiento** [3, 17, 18]: a partir de conocimiento sobre los usuarios y los productos se persigue un razonamiento que indique qué producto cumple los requerimientos del usuario, dejando un poco a un lado valoraciones que el usuario pueda hacer.
- **Sistemas de recomendación basados en utilidad** [6, 19]: estos sistemas realizan las recomendaciones basándose en el cálculo de la

utilidad de cada objeto en particular con respecto al perfil determinado de cada usuario.

- **Sistemas de recomendación híbridos** [12, 14, 19-24]: este tipo de recomendadores surgió con el objetivo de solventar algunos problemas presentados por los sistemas anteriores ante algunas situaciones. Para ello se realizan combinaciones entre dos o varias de las diferentes técnicas de funcionamiento anteriores.

Para terminar esta breve introducción a los sistemas de recomendación es importante hacer notar que estos sistemas, sean del tipo que sean, utilizan casi exclusivamente dos tipos de información:

- **Información explícita** [2, 5, 9, 12, 25-29]: información proporcionada por el usuario, en la mayoría de los casos a petición del sistema, aunque también puede ser introducida por terceras personas en base a cuestionarios rellenados, datos aportados, información ya disponible, etc., pero siempre contemplando información directamente proporcionada por el usuario; de esta forma es dicho usuario el responsable de la veracidad de la información aportada.
- **Información implícita** [2, 5, 9, 12, 25, 26, 28-30]: información recogida automáticamente por el propio sistema en función del comportamiento del usuario. Pueden ser por ejemplo datos referidos al historial de navegación del usuario, productos adquiridos en una página web, acceso a ciertas páginas concretas, etc.

Es basándose en esta información como los sistemas de recomendación consiguen realizar sus predicciones, y en la mayoría de los casos, el tipo de información con la que se pretende trabajar, o simplemente la información de la que se dispone, puede delimitar o definir las técnicas de recomendación que se pueden aplicar.

1.2 Motivación de este trabajo

Aunque muchos campos y ámbitos se han aprovechado del positivo influjo proporcionado por los sistemas de recomendación, existen todavía áreas poco exploradas por estas tecnologías en las que el uso de estas herramientas podría resultar enormemente útil. Uno de estos campos en los que queda aún mucho camino por recorrer es el que concierne a la educación.

Podemos decir que los cambios que se han producido en la educación, sobre todo en los sectores más reglados y elementales, se resumen prácticamente en sustituir la tiza por un puntero, la pizarra por un proyector, y las enciclopedias por Internet, sin ni siquiera haber tenido en cuenta muchas veces los peligros que implica un uso desatinado de esta fuente de información; y digo esto desde el punto de vista del profesor de educación secundaria que soy. Ciertamente es innegable que ha habido numerosas mejoras relevantes en éste sector

provenientes de las nuevas tecnologías, pero están ceñidas sobre todo a aquellas enseñanzas de carácter post-obligatorio y/o profesional; es más, se podría decir que la mayoría de dichas mejoras suelen quedar restringidas a la creación y mejora de plataformas para la gestión de contenidos y/o material didáctico o para la enseñanza a distancia mediante el uso de ordenadores (*e-learning*, o enseñanza *online*).

Como ejemplo de esto último que acabamos de decir, existen algunos intentos de proponer el uso de sistemas de recomendación colaborativos en el ámbito de la educación *online*, como por ejemplo en [31], en el que se argumenta que estas técnicas facilitarían el aprendizaje colaborativo e independizarían aún más al alumnado usuario de este tipo de enseñanza, sin embargo no se propone nada concreto. En [32] se plantea el uso de sistemas de recomendación a la hora de elegir el *CourseWare* que mejor se ajuste al entorno educativo que se quiere aplicar, y se integra este recomendador en una herramienta de mantenimiento de plataformas *CourseWare*. Otro acercamiento de los sistemas de recomendación a la educación *online* presenta el objetivo de realizar recomendaciones personalizadas sobre contenidos de cursos (concretamente sobre objetos de aprendizaje), para lo cual hace uso de estrategias basadas en algoritmos colaborativos mezcladas con recuperación de información [30]. Por último, y otra vez en el ámbito de la educación a distancia, se sugiere un sistema de recomendación de recursos de aprendizaje y de acercamiento de personas con intereses parecidos a la hora de aprender basado en sistemas colaborativos [29].

Es así, a pesar de estos intentos y tras varios años en la docencia, he comprobado que en la práctica existen realmente escasas opciones en el ámbito de la educación elemental que se aprovechen de los avances realizados en el campo de los sistemas inteligentes. En el mejor de los casos debemos contentarnos con utilizar herramientas cuyo fin fue inicialmente otro, y que son adaptadas (o más bien nosotros nos adaptamos a ellas) a la hora de utilizarlas en el ámbito de la enseñanza. Ejemplo bastante significativo de esto es el uso de plataformas de enseñanza *online* en aulas de enseñanza presencial, hecho tan extendido que hasta se le ha dado nombre: *blended-learning* [33, 34]. No es improbable encontrar profesores que de manera quijotesca se proponen el mejorar su docencia ayudándose de plataformas tales como MOODLE, que incluso con sus carencias en términos de docencia presencial, aportan beneficios significativos a la labor del profesorado y mejoran considerablemente el rendimiento del alumnado.

Por todo esto, lo que aquí se propone es dar un paso que acerque los sistemas de recomendación al alumnado, por supuesto con el fin de ayudarles en algún tipo de tarea, pero también con el propósito de que conozcan y se familiaricen con nuevas herramientas y nuevas vías para conseguir aquellos objetivos que en un momento dado puedan perseguir.

Ahora bien, no es una tarea fácil el decidir dónde actuar. Encontramos a este respecto diversos frentes de actuación, en algunos de los cuales ya se está trabajando: avances centrados en mejorar las plataformas educativas existentes, proporcionándoles por ejemplo cierto grado de adaptatividad en función del alumno en cuestión, decidiendo qué tareas son buenas para él y cuáles no, qué contenidos mostrarle y cómo, adaptando el sistema de evaluación, la interfaz en función de cada individuo, etc.

Sea como fuere, el objetivo que cualquiera de estos desarrollos persiguen es claro: facilitar el aprendizaje del alumno mejorando la forma en la que se le presentan las actividades y contenidos con los que debe trabajar. Esto se puede conseguir tanto actuando directamente sobre dicho aprendizaje con aportaciones como las expuestas en el párrafo anterior, como facilitando la labor profesorado a la hora de atender a la gran cantidad de alumnos que maneja.

Sin embargo podemos plantearnos otro campo en el que tomar partido; existe hoy día una figura propia de la educación preuniversitaria, una figura que sirve de nexo entre educadores y estudiantes y que cada vez cobra más importancia: el *orientador*. Muchas son las funciones asignadas a este cargo: desde mediar en conflictos, pasando por proporcionar apoyo emocional, hasta cumplir la tarea que le da nombre, orientar a los miembros de la comunidad educativa. Y sobre todo orientar al alumnado a la hora de tomar ciertas decisiones. Decisiones del tipo: *Para el futuro que deseo, ¿qué tipo de estudios debo elegir?, ¿a qué área me debo orientar?, ¿qué asignaturas debería cursar para hacer esto realidad?*

Aquí es donde vamos a centrar nuestro estudio, analizando la posibilidad de utilizar un sistema de recomendación automatizado encargado de proporcionar ayuda y orientar en la medida de lo posible al alumnado y al personal de orientación a la hora de determinar los posibles caminos formativos a elegir en torno a las distintas posibilidades educativas que pueden presentarse.

El fundamento teórico sobre la orientación educativa y las normativas legales aplicadas a la Educación Secundaria Obligatoria y al Bachillerato pueden encontrarse en [35-40]. En el punto 3 hablaremos ampliamente del dominio concreto en el que nos moveremos y se hará una breve introducción a estas etapas educativas.

1.3 Propósito

La idea que se persigue en esta investigación es comprobar hasta qué punto un sistema de recomendación, en concreto, un sistema de recomendación basado en filtrado colaborativo (en adelante CF – *collaborative filtering*), es capaz de, en función de la trayectoria académica que un alumno o alumna ha seguido, predecir las asignaturas en las que el alumno podría obtener mejores resultados, y analizar la posibilidad de recomendar aquella asignatura o aquel

grupo de asignaturas para las que ha demostrado tener más capacidad o le podrían resultar de más interés.

Se ha elegido utilizar técnicas basadas en filtrado colaborativo por diversas razones [5, 8]:

- Su sencillez y mayor intuitividad con respecto a otras técnicas más complejas y de similares resultados;
- su capacidad para tratar con elementos difíciles de analizar mediante procesos informáticos;
- el tipo de información con la que estos sistemas manejan se adapta de forma simple y directa a la información con la que nosotros pretendemos trabajar, es decir, las calificaciones del alumnado disponibles en los Institutos de Educación Secundaria (IES);
- permiten seleccionar ítems basándose en medidas tales como *calidad, adecuación* y *gusto*;
- presentan una interesante capacidad a la hora de proporcionar unas recomendaciones personalizadas de calidad;
- reducen la sobrecarga de información en los distintos ámbitos en los que se han utilizado, delimitando de una forma clara y sencilla las distintas alternativas o vías de actuación;
- en ocasiones proporcionan recomendaciones que para el usuario resultan novedosas e interesantes;
- y para terminar, para éste dominio en concreto, como veremos más adelante, los problemas presentados de forma inherente por los sistemas de recomendación basados en filtrado colaborativo se reducen de forma considerable.

Es muy importante tener en cuenta que lo que aquí se intenta es evaluar el comportamiento de estas técnicas a la hora de estimar puntuaciones, que es lo que hemos visto que realmente hacen los sistemas de recomendación antes de producir un resultado concreto, pero sin perder de vista el alcance de esta predicción y el carácter meramente orientativo de la recomendación. Por muy bien que se realice, por muy certera que pueda parecer, estamos hablando de seres humanos, de adolescentes, y en el caso de que los resultados resultaran satisfactorios habría que realizar un estudio sobre cómo adecuar esta valoración de tipo *cuantitativa*, adaptándola de forma que tomara en cuenta otros factores más *cualitativos* referidos a las actitudes del alumnado, sus aptitudes, orientaciones académicas, así como sus preferencias.

Todo esto será debatido posteriormente en la sección correspondiente a *Trabajo Futuro*.

1.4 Aportaciones

- Se ofrece una revisión bastante general del estado del arte en lo referente a sistemas de recomendación basados en filtrado colaborativo hasta la fecha, centrada sobre todo en lo concerniente a algoritmos basados en memoria (Capítulo 2).
- Se estudian las extensiones y mejoras que para este tipo de algoritmos han aparecido hasta el momento (Capítulo 2, Epígrafe 2.5).
- Se estudia el impacto que sobre este tipo de ítems puede llegar a ejercer el filtrado colaborativo, teniendo en cuenta las características propias del dominio, alejadas de las de aquellos ítems objetivo de las aplicaciones realizadas hasta la fecha: valoraciones que no expresan gusto explícito del usuario, sino su comportamiento objetivo frente a un ítem determinado (Capítulo 3).
- Se evalúa el uso de estas técnicas a la hora de recomendar objetos que a priori no parecerían susceptibles de utilizar con este tipo de algoritmos, refiriéndonos claro está a las asignaturas (Capítulo 4).
- Se propone una configuración de parámetros optimizada para un algoritmo que se desenvuelva en este ámbito de trabajo atendiendo a los datos de los que se disponen (Capítulo 4, Epígrafe 4.8).
- También se evalúa el impacto de estas técnicas en datos que no son explícitos, pero que tampoco pueden considerarse implícitos, por lo menos a la manera en la que se viene haciendo en la amplia mayoría de los sistemas desarrollados, puesto que las valoraciones ni las introduce el usuario, ni las calcula automáticamente el sistema en base al comportamiento de dicho usuario. Las calificaciones las introducen expertos después de evaluar el comportamiento del alumno frente a una serie de objetivos a cumplir (Capítulos 4 y 5).
- Se elabora un sistema de recomendación denominado OrieB para orientar al alumnado a la hora de enfrentarse al Bachillerato (Capítulo 6).
- Por último, se aportarán unas conclusiones derivadas de estos estudios y evaluaciones y en caso de que los resultados resulten prometedores, se propondrán diversas alternativas y vías de actuación a la hora de plantear y desarrollar un sistema real (Capítulo 7).

2. SISTEMAS DE RECOMENDACIÓN BASADOS EN FILTRADO COLABORATIVO

2 SISTEMAS DE RECOMENDACIÓN BASADOS EN FILTRADO COLABORATIVO

La totalidad de la literatura existente está de acuerdo en que los sistemas basados en filtrado colaborativo trabajan recogiendo juicios humanos, expresados como votaciones de cada individuo para una serie de ítems en un domino dado, y tratan de emparejar personas que comparten las mismas necesidades o gustos [5, 8-10, 12, 13, 41-43].

Los usuarios de un sistema colaborativo comparten sus valoraciones y opiniones con respecto a los ítems que conocen de forma que otros usuarios del sistema puedan decidir qué elección realizar. A cambio, el sistema proporciona recomendaciones personalizadas para aquellos elementos que resultan interesantes para el usuario.

Estos sistemas proporcionan diversas ventajas con respecto a otros sistemas de recomendación como acabamos de ver en el capítulo anterior, y han sido probados con éxito en diversos dominios, sobre todo en aquellos relativos al entretenimiento [43-46].

Es importante tener en cuenta que en el filtrado colaborativo son los usuarios, las personas, quienes determinan la relevancia, cualidad e interés de los ítems, por lo que se puede realizar el filtrado sobre elementos difíciles de analizar mediante computación. Esto se consigue gracias a que el problema no se intenta solucionar analizando los elementos que tratamos de recomendar, sino manejando directamente las valoraciones, objetivas y/o subjetivas, que sobre esos elementos se han realizado.

Además, el filtrado colaborativo tiene la capacidad de discernir cómo se adapta un ítem a las necesidades o intereses de los usuarios, basándose en la propia capacidad de los humanos de analizar en términos de calidad o gusto, algo difícil de realizar por procesos computacionales.

Por último, un sistema colaborativo puede aportar al usuario novedad, haciendo recomendaciones de ítems valiosos que dicho usuario no esperaba encontrar, y que de otro modo posiblemente jamás habría considerado. El término para este suceso en inglés es *serendipity*, y hace referencia a un hallazgo casual beneficioso, afortunado y no esperado [2, 9, 47].

Como vemos, el potencial del filtrado colaborativo es enorme a la hora de discriminar y tener en cuenta comportamientos subjetivos. Sin embargo, según [8], para alcanzar el potencial completo, es casi seguro que debe combinarse con tecnologías de sistemas basados en contenido.

Los algoritmos de recomendación colaborativos pueden ser agrupados en dos clases generales [2, 5, 27, 41, 48-53]:

- **Algoritmos basados en memoria** (o en vecindad, o heurísticos)
- **Algoritmos basados en modelos**

Los algoritmos basados en memoria son esencialmente heurísticas que realizan predicciones basadas en una colección completa de ítems valorados previamente por el usuario. Es decir, el valor de una puntuación no conocida *v* para un usuario *u* sobre un ítem *i* se calcula como un agregado de las valoraciones de otros usuarios (generalmente, los *K* más parecidos) para el mismo ítem *i*.

En contraste con esto, las aproximaciones basadas en modelos utilizan la colección de valoraciones para aprender un *modelo*, el cual será utilizado a la hora de realizar las futuras predicciones. Algunos algoritmos colaborativos *basados en modelos* se fundamentan en:

- Uso de redes bayesianas [41] y derivados [54].
- Métodos de *clustering* para filtrado colaborativo [11, 13, 53].
- Análisis iterativo de la componente principal [10].
- Otro método interesante es utilizar modelos de aprendizaje de forma que se pueda tratar el filtrado colaborativo como un *problema de clasificación* [7].

Desde este momento vamos a centrarnos en los algoritmos basados en memoria o vecindad, debido a su mejor adecuación a nuestro problema.

A continuación vamos a revisar una serie de conceptos que han de conocerse sobre los Sistemas de Recomendación Colaborativos.

2.1 Dimensión del problema

Según [9], podemos distinguir diversas tareas que pueden plantearse como objetivo de un sistema:

- **Anotación en contexto**: en dominios específicos, se trata de determinar si para el usuario merece la pena tomarse el tiempo de examinar o no un ítem. Por ejemplo, podemos usar el filtrado colaborativo para determinar qué artículos dentro de un área merece la pena leer, qué película puede resultar interesante ver, qué música podría gustarnos oír, etc.
- **Encontrar ítems buenos**: se trata de sugerir al usuario una serie de objetos proporcionando una lista ordenada y/o puntuada de ítems recomendados.

	Dire Straits	Metallica	The Beatles	Depeche Mode	Bee Gees
Alex	4	1	9	6	10
Ricardo	7	9	6	10	
Eva	5	2	9	7	8
Pedro	6	3	?	6	7

Tabla 1: Representación del espacio de un problema asociado a filtrado colaborativo.

Si la información que estamos manejando consta de una serie de ítems, una serie de usuarios, y una serie de valoraciones de estos usuarios sobre aquellos ítems, podemos concluir que el espacio del problema viene definido como una matriz de usuarios frente a ítems, en la que cada celda representa la puntuación de un usuario referida a un ítem específico [48] (Tabla 1).

En base a esto, resolver el problema bien para tareas de anotación en contexto, bien para encontrar buenos ítems, implica predecir el grado en el que un ítem gustará a un usuario que no tiene puntuación asociada a ese ítem, mediante el uso de una serie de valoraciones aportadas anteriormente por un grupo de usuarios: predecir los valores para aquellas celdas que se encuentran vacías [5, 8], o lo que es lo mismo, estimar qué valoración daría el usuario a aquellos ítems que todavía no ha valorado (Figura 1).

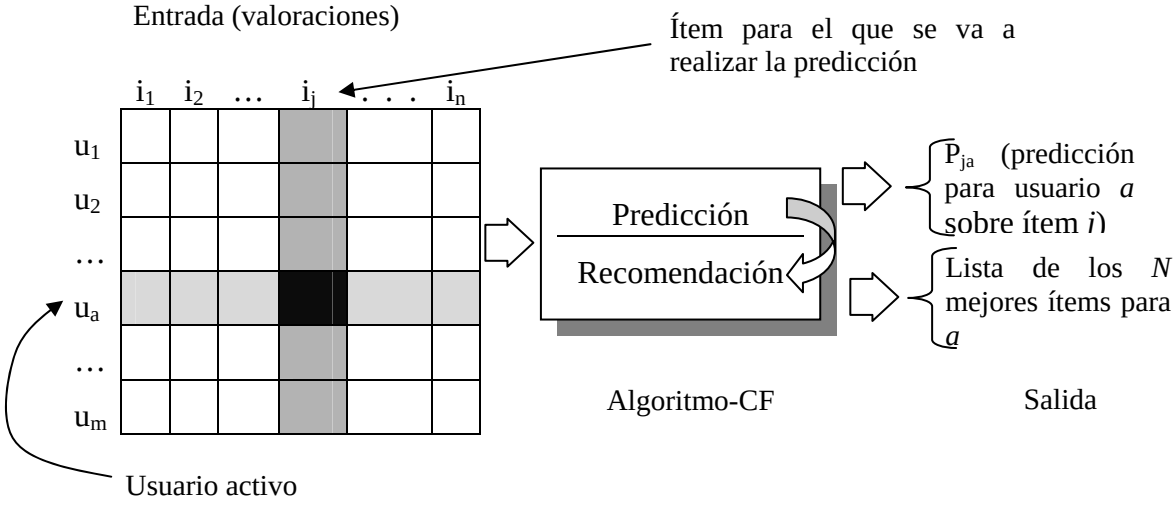


Figura 1. Formación de predicciones en base a las valoraciones de los usuarios.

Los resultados mostrados al usuario se pueden presentar como una simple predicción, o como una recomendación formada por una lista de los *N* ítems que al usuario deberían gustarle más, teniendo para ello en cuenta si se deben mostrar ítems anteriormente votados o no.

2.2 Análisis de los datos

Existen diversas cuestiones a tener en cuenta cuando nos enfrentamos al análisis de los datos con los que vamos a trabajar en un sistema de filtrado colaborativo [9], y que pueden determinar ciertas decisiones en pasos posteriores. Es posible que algunas veces tengamos que elegir una de entre varias alternativas, sacrificando el resto, o bien intentar hacer que convivan tales opciones, teniendo que tomar estas decisiones en cuenta para implementaciones posteriores.

2.2.1 Características del dominio

Habría muchas a tener en cuenta, pero nos vamos a centrar en las más importantes [9]:

- El trasfondo y contexto del ítem a recomendar
- La tarea que pretende resolver el sistema
- La novedad que se pretende que aporte la recomendación
- La calidad necesitada
- La proporción entre coste y beneficio, de los falsos/verdaderos positivos/negativos
- La granularidad de las preferencias reales del usuario

Primero deberemos tener en cuenta el **trasfondo** del elemento a recomendar y el **contexto** en el que la recomendación tomará lugar, puesto que no es lo mismo aconsejar música o libros que proponer la compra de un coche, o aconsejar a la hora de tomar una decisión relevante. Otro factor a tener en cuenta y que tiene que ver con esto es la **tarea objetivo** a la que pretende dar solución el sistema; en función de un tipo de tarea o de otra la metodología de trabajo puede variar.

Es muy importante reflexionar sobre qué debemos primar, si la **novedad** de la recomendación que se va a realizar, o la **calidad** de esa recomendación. Generalmente no es sencillo, a veces ni siquiera posible, optimizar ambos factores a la vez. Si por ejemplo tuviéramos que recomendar un libro, primar la novedad implicaría intentar buscar un hallazgo que sorprenda al usuario, mientras que si simplemente buscamos la calidad, el sistema podría recomendar *valores seguros* (como los 10 más vendidos, o el libro del mes, por ejemplo), que dependiendo del caso pueden no satisfacer las necesidades que nos habíamos propuesto cubrir. Esto cambia si lo que prevalece es la necesidad de una alta

confianza con respecto a las predicciones dadas, y lo que se busca es que las decisiones sean siempre realizadas con atino.

Un **falso positivo** es una recomendación no acertada, un elemento que se ha recomendado pero que al usuario no le gusta, mientras que un **falso negativo** es no recomendar (o simplemente dejar de recomendar) un elemento que debería haberse recomendado [48, 55], o según el caso, ofrecer una recomendación negativa sobre un objeto, que nos haga pensar que dicho objeto no es digno de consideración. .

Sabiendo esto, es muy importante estudiar el **factor coste/beneficio** de estos falsos negativos y positivos. Un libro mal recomendado (falso positivo) tendrá como coste el precio del libro (y algunos opinarían que el tiempo para leerlo); un libro no recomendado (falso negativo) realmente puede considerarse como coste cero; un libro bien recomendado en este caso proporciona un beneficio considerable de cara al usuario. En torno a estas posibilidades, habría que estudiar hasta que punto se pueden tolerar los falsos positivos y los falsos negativos.

En la mayoría de los sistemas de recomendación el tipo más importante de error a evitar es el que concierne a los falsos positivos: un falso negativo generalmente es difícil de detectar por un usuario, que no podría no llegar nunca a saber de la existencia de un producto; sin embargo, un falso positivo puede desencadenar en un enfado del usuario, y de aquí en una falta de confianza con respecto al sistema de recomendación en cuestión.

Por último, es necesario tener en cuenta la **granularidad** de las preferencias reales del usuario, los distintos niveles en los que estas preferencias se pueden clasificar: pueden considerarse binarias (con dos niveles, expresando si al usuario le gusta o no), diferentes tipos de escalas (de 1 a 5 o de 1 a 10 por ejemplo), escalas con valores difusos [6], etc.

2.2.2 Características inherentes

Incluyen diversas propiedades referidas a las valoraciones:

- Valoraciones implícitas, explícitas, o ambas.
- Escala en la que las puntuaciones se realizarán.
- Dimensión de las valoraciones.
- Presencia o ausencia de marca temporal en las valoraciones.

Las **valoraciones explícitas** son generalmente puntuaciones numéricas aportadas directamente por el usuario en las que un número alto indica un alto grado de interés por un ítem, y puntuaciones bajas expresan desinterés. Para obtener estos juicios generalmente se alecciona a los usuarios a puntuar los ítems de la forma en la que a ellos les hubiera gustado haber visto valorado tal objeto.

Lógicamente, esto requiere cierto esfuerzo y tiempo por parte del usuario al ser él el encargado de proporcionar al sistema la información requerida.

Las **valoraciones implícitas** pueden ser obtenidas de diversas fuentes de datos tales como registros de compras, historial de navegación, etc., tareas que al usuario no se solicita realizar de una forma manifiesta, es decir, se infieren automáticamente por el sistema de recomendación a partir del comportamiento del usuario, por lo que no se requiere la atención del usuario para su recolección.

La **escala de las puntuaciones** tiene que ver con la granularidad de las valoraciones. La más simple es unaria, se marcan los ítems que gustan y el resto se desconoce su valoración (no se puede decir que no gusten, simplemente se desconoce este hecho). Por ejemplo, puede tratarse de los objetos comprados por un cliente. Valores binarios implican representación para los elementos que no gustan. Existen escalas para valoraciones explícitas usadas generalmente de 1 a 5, de 1 a 7, de 1 a 10, e incluso de 1 a 100, También pueden proporcionarse valoraciones múltiples, puntuando diversas características del producto (podrían ser por ejemplo en el caso de vehículos: apariencia, fiabilidad, potencia, confort, etc.). Veremos esta cuestión más adelante cuando hablemos de la *granularidad de las valoraciones* en el punto *3.2.1,*

2.3 Pasos o tareas de un algoritmo de filtrado colaborativo basado en memoria

Para los métodos basados en memoria, y una vez definidos los datos con los que se debe trabajar, las tareas a realizar en base a generar una predicción son tres [5], [8] (Figura 2):

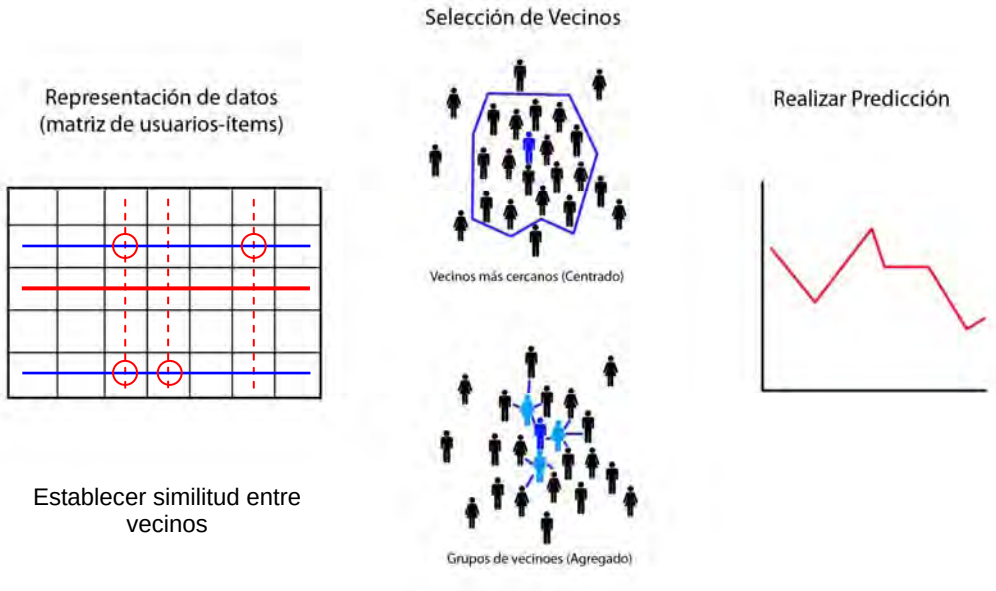


Figura 2. Las tres partes principales de un sistema de recomendación CF

1. Establecer el valor de similitud entre el usuario activo y el resto de usuarios.
2. Seleccionar un conjunto de usuarios que se usarán para la predicción.
3. Normalizar las valoraciones y generar una predicción.

2.3.1 Establecer similitud entre vecinos

Para establecer la similitud entre los vecinos de los que obtendremos la predicción debemos definir una medida que nos permita evaluar el grado de parecido entre unos y otros.

Medidas de similitud

En [2, 5, 8, 12, 41, 48, 52] podemos encontrar las medidas de similitud generalmente utilizadas para evaluar la similitud entre usuarios. De entre todas, nosotros vamos a revisar las más referenciadas en la literatura sobre sistemas de recomendación. Estas medidas son:

Coeficiente de Correlación de Pearson

El Coeficiente de Correlación de Pearson (en adelante PCC – *Pearson correlation coefficient*) es la primera formulación estadística aparecida para el filtrado colaborativo. La correlación entre el usuario *u* y el usuario *i* se define como:

$$w(u,i) = \frac{\sum_j (v_{u,j} - \overline{v_u})(v_{i,j} - \overline{v_i})}{\sqrt{\sum_j (v_{u,j} - \overline{v_u})^2 \sum_j (v_{i,j} - \overline{v_i})^2}}$$

(Ecuación 1, PCC)

Vector de Similitud o Distancia del Coseno

En el campo de la recuperación de información [56, 57], la similitud entre dos documentos se mide a menudo tratando cada documento como un vector de frecuencias de palabras y calculando el coseno del ángulo formado por los dos vectores de frecuencia. Podemos adoptar este formalismo al filtrado colaborativo tomando los usuarios el rol de documentos, los ítems el rol de palabras y las valoraciones el rol de frecuencias. Es importante destacar que los votos deben ser positivos para ser tenidos en cuenta. En base a [5, 41] se puede formular esta medida de similitud (desde ahora COS – *Cosine*) como sigue:

$$w(u,i) = \cos(\vec{v}_u, \vec{v}_i) = \frac{\vec{v}_u \vec{v}_i}{\left\|\vec{v}_u\right\|^2 \times \left\|\vec{v}_i\right\|^2} = \frac{\sum_j v_{u,j} v_{i,j}}{\sqrt{\sum_j (v_{u,j}{}^2)} \sqrt{\sum_j (v_{i,j}{}^2)}}$$

(Ecuación 2, COS)

Estas dos medidas definen un recurso que nos permitirá elegir una serie de usuarios con los cuales poder realizar una predicción. Sin embargo, existen extensiones que permiten en cierto modo mejorar la precisión de estas medidas, como veremos más adelante.

2.3.2 Selección de vecinos

Una vez que sabemos cómo calcular la similitud entre un usuario concreto, el usuario activo, y el resto de usuarios (Figura 2), se hace necesario elegir cuáles de entre esos usuarios se utilizarán para computar las predicciones. En términos de eficiencia y exactitud se debe elegir un subconjunto de usuarios para el cálculo de la predicción en vez de usar el conjunto completo. Este problema puede entenderse como un problema típico de clasificación, en el que se trata de agrupar a los usuarios en grupos o clases: los que nos van a servir para elaborar la predicción, y los que no son considerables.

En [8] y [42] se evalúa la utilización de dos posibles técnicas para determinar cuántos vecinos utilizar. Estas técnicas son la utilización de un umbral y la elección de los mejores *k* vecinos, conocida como *vecindad centrada* (en adelante KNN – *k nearest neighbors*).

En la primera se propone un valor umbral, eligiendo todos aquellos vecinos que lo superen. Esto puede presentar algunos problemas, como por ejemplo que no haya un número suficiente de vecinos que supere dicho umbral para umbrales elevados, provocando problemas de cobertura, o bien que el número de vecinos elegidos sea excesivamente grande, computacionalmente hablando, debido a un umbral bajo, haciendo inútil el intentar elegir un subconjunto.

La segunda técnica consiste en elegir un número *K* y utilizar los *K* vecinos con mayor factor de similitud, evitando que haya o pocos o muchos usuarios. La elección de éste número *K* es crucial puesto que si elegimos un número demasiado grande habrá un ruido considerable en las predicciones, perjudicando la exactitud de las mismas, pero si el número es demasiado pequeño, la cobertura disminuirá y habrá muchas predicciones que no podrán realizarse.

Combinar ambas técnicas no mejora mucho el uso de KNN por lo que la esencia está en encontrar un número adecuado para *k*.

Existe otro tipo de técnica a la hora de seleccionar vecinos para casos en los que la dispersión es acuciada [48]. Se trata de un *agregado de vecinos* en el que se utiliza el vecino más parecido al usuario activo, el resto de los *K*–1 vecinos se calcula en base al cálculo del centroide de la vecindad, lo que permite a los vecinos más cercanos tomar parte en la formación de dicha vecindad, lo que puede ser beneficioso para conjuntos de datos dispersos.

2.3.3 Realización de predicciones

Una vez que se ha escogido un vecindario, debemos combinar las valoraciones de ese vecindario para producir una predicción [5, 12, 48, 58]. La forma más sencilla de combinar estas puntuaciones es calcular una media con las valoraciones, sin embargo una de las aproximaciones más utilizadas es calcular una suma media ajustada de las puntuaciones, utilizando las correlaciones como pesos. Esta suma ponderada (en adelante WS – *weighted sum*) supone que todos los usuarios puntúan siguiendo aproximadamente la misma distribución. Se suele utilizar un multiplicador *m* como factor de normalización, y generalmente se utiliza $m=1/\Sigma_i |w_{u,i}|$, es decir, el inverso de la suma de correlaciones. Además de la suma ponderada se puede utilizar también una suma media ajustada (en adelante WA – *weighted adjusted sum*) que pretende evitar un problema que presenta WS, y que consiste no tomar en consideración el hecho de que usuarios diferentes puedan usar escalas diferentes de puntuación, es decir, diferentes usuarios pueden puntuar siempre o muy alto, o muy bajo, o siempre en los extremos, o sólo términos medios. Para ello en WA se utilizan desviaciones de la media para el usuario correspondiente.

$$v_{u,i} = \frac{\sum_j w_{u,j} v_{j,i}}{\sum_j w_{u,j}}$$

(Ecuación 3, WS)

$$v_{u,i} = \overline{v_u} + \frac{\sum_j w_{u,j} (v_{j,i} - \overline{v_j})}{\sum_j w_{u,j}}$$

(Ecuación 4, WA)

Así podremos obtener predicciones para los ítems que deseemos y que el usuario no haya puntuado.

Según [8], el usuario activo proporciona al motor de predicción una serie de ítems para los que desea obtener predicciones y el motor de predicción le devuelve una lista de predicciones para esos ítems. En otros casos, el motor, tras realizar los cálculos, lo que de vuelve es una lista puntuada de las *n* mejores predicciones. Es imprescindible para proporcionar una interfaz fluida el establecer unas restricciones de eficiencia, dado que una respuesta de una latencia de 1 o más segundos no es considerable, y un motor se enfrenta a cientos de solicitudes por segundo.

2.4 Problemas presentados por los algoritmos colaborativos basados en el usuario

Hasta ahora, la base de los sistemas de recomendación basados en filtrado colaborativo ha sido centrarse en buscar similitudes en las preferencias de los distintos usuarios y proporcionar una predicción sobre los distintos ítems en función de las calificaciones que aquellos usuarios con gustos parecidos al usuario objetivo realizaron en su momento. Todo esto parte de que la idea de que es bueno recomendar a una persona ítems que funcionaron en el pasado con usuarios de gustos similares.

Este tipo de algoritmos, denominados algoritmos basados en el usuario (a partir de ahora CF-U), han tenido mucho éxito, pero su uso enormemente extendido ha revelado algunos problemas que implican retos potenciales [48]. Esto quiere decir que dichos sistemas presentan de forma inherente algunos problemas que vienen dados la forma misma en que el problema se formula.

Pasamos a explicar estos posibles problemas:

- **Escalabilidad**
  Para que el sistema pueda realizar predicciones de una exactitud aceptable, es completamente necesario tener un número suficiente de usuarios que han votado ítems. Esto implica una doble vía de crecimiento: el sistema hará predicciones más exactas conforme los usuarios voten más ítems, y también conforme crece el número de usuarios que voten ítems. En las aplicaciones comerciales, esto implica que el número de ternas *usuario-ítem-votación* crece enormemente creando problemas de escalabilidad, de forma que la computación que se requiere para calcular la vecindad más cercana crece con el número de usuarios, con el número de ítems y con el número de votaciones realizadas, pudiendo sufrirse grandes penalizaciones temporales [22, 27, 43, 50, 51, 58].

- **Nuevo-usuario**
  Los usuarios nuevos se ven forzados a calificar un número suficiente de ítems antes de que el recomendador pueda comprender las preferencias del usuario y presentar sugerencias de confianza [5, 9, 10, 26, 43].

- **Nuevo-ítem**
  Si se añaden al sistema nuevos ítems, no podrán ser recomendados hasta que sean votados por un número sustancial de usuarios [5].

- **Dispersión**
  En dominios en los que existe una gran cantidad de ítems a valorar es difícil encontrar usuarios que hayan votado elementos similares, de forma que se requiere una masiva cantidad de usuarios para disminuir la probabilidad de que esto ocurra. También es posible que ítems con

gran potencial de aceptación pero escasamente votados sean poco recomendados porque no se llegue a ellos [5, 9, 12, 22, 32, 50, 51].

- **Ítems-sinónimos**
  En [5, 48] se plantea el problema de los **ítems sinónimos**, es decir, que existan productos que con distintos nombres hagan referencia a objetos similares (desde el punto de vista colaborativo), o dicho de otro modo, ítems que con distinto nombre tengan iguales propiedades (desde el punto de vista basado en contenido). El sistema puede no llegar a contemplar esta relación y tratar de forma diferente dichos objetos.

En la práctica, muchos sistemas de recomendación comerciales evalúan grandes conjuntos de ítems, y esto puede provocar que se presenten de una forma acuciante problemas de escalabilidad, y peor aún, de dispersión, pudiendo darse el caso de que para un usuario concreto no pueda darse ninguna recomendación puesto que los ítems que ha votado no son suficientes para establecer una vecindad, o bien puede degradarse la precisión de las predicciones debido a la propia dispersión de puntuaciones.

El problema de los ítems sinónimos es más acusado en los métodos basados en contenido, puesto que presenta la desventaja añadida de que dos elementos tomados como similares debido a sus características compartidas, pueden considerarse de igual calidad; esto es como decir que dos coches azules, de 5 puertas, y 140 CV tienen la misma calidad o cumplen las mismas funciones. Sin entrar en cuestiones comerciales, sabemos que esto no es cierto.

En los algoritmos de filtrado colaborativo, el problema de los ítems sinónimos afecta sobre todo a la hora del cálculo de similitud entre vecinos, pudiendo pasarse por alto gustos similares al denominarse de forma distinta determinados productos.

2.5 Mejoras para los Algoritmos de Filtrado Colaborativo

En [5, 8, 41] se definen una serie de posibles extensiones aplicables a las medidas de similitud y técnicas de predicción antes vistas y que dependiendo del caso pueden mejorar la exactitud del sistema.

También se hará mención de una serie de mejoras aplicables en otras fases de la resolución del problema.

Frecuencia inversa

La idea que esta mejora persigue es reducir el peso de aquellos ítems que son votados por la gran mayoría de la comunidad, ya que por intuición no resultarán discriminatorios a la hora de diferenciar usuarios, mientras que aquellos ítems que se votan menos toman un mayor peso. Se define la frecuencia inversa como [41]:

$$f_j = \log \frac{n}{n_j} \quad \text{(Ecuación 5)}$$

donde n_j es el número de usuarios que han votado por el ítem *j*, y *n* es el número total de ítems en la base de datos. Hay que tener en cuenta que si todo el mundo ha votado un ítem *j*, entonces f_j será cero.

El cálculo de la similitud con COS teniendo en cuenta la frecuencia inversa se realizaría simplemente multiplicando el voto original por f_j: Sin embargo, para PCC se debe modificar la Ecuación 2 quedando como sigue:

$$w(u,i) = \frac{\sum_j f_j \sum_j f_j v_{u,j} v_{i,j} - \left(\sum_j f_j v_{u,j} \right) \left(\sum_j f_j v_{i,j} \right)}{\sqrt{\sum_j f_j \left(\sum_j f_j v^2{}_{u,j} - \left(\sum_j f_j v_{u,j} \right)^2 \right)} \sqrt{\sum_j f_j \left(\sum_j f_j v^2{}_{i,j} - \left(\sum_j f_j v_{i,j} \right)^2 \right)}}$$

(Ecuación 6)

Amplificación de casos

Se refiere a la transformación aplicada a los pesos usados en la fórmula de las predicciones, de forma que se persigue enfatizar las similitudes mayores, es decir, los pesos más cercanos a 1, penalizando los más lejanos [41]. La constante ρ suele tomar como valor 2.5 para los experimentos.

$$w'_{u,i} = \begin{cases} w^{\rho}{}_{u,i} & si \quad w^{\rho}{}_{u,i} \geq 0 \\ -(-w_{u,i})^{\rho} & si \quad w^{\rho}{}_{u,i} < 0 \end{cases} \quad \text{(Ecuación 7)}$$

Voto por defecto

El voto por defecto es una extensión de PCC (Ecuación 2) y que se basa en la observación de que si nos encontramos con usuarios que tienen relativamente pocos votos, ya sea el usuario activo o aquél con el que queremos medir la similitud, el algoritmo de correlación no funciona bien debido a que usa

sólo los votos de la intersección de los ítems que ambos usuarios han votado [41].

Así, esta alternativa propone asumir un voto por defecto para un número de ítems adicionales que los usuarios no han votado.

Factor de relevancia

Como en [12] se hace notar, las medidas de similitud son más significativas cuando hay muchos ítems votados en común entre los usuarios. Se introduce así en [8] un factor para ponderar la importancia de la similitud entre dos usuarios. La idea es proporcionar una *confianza* a la similitud entre usuarios. Para entender esto baste decir que no es muy fiable el que dos usuarios tenga una similitud de 1 basándose únicamente en una votación común, y que similitudes menores pero basadas en gran número de votaciones comunes pierdan relevancia frente a esta.

Cuantos más ítems comunes se estén comparando, mayor y más representativa debería ser la fiabilidad de esa medida de similitud. Se propone aplicar un parámetro que exprese la relevancia de una medida de similitud entre usuarios, aportando confianza a esta medida. La nueva medida de relevancia quedaría como sigue:

$$w'_{u,i} = w_{u,i} \cdot \frac{n}{N}, \qquad si \quad n < N \qquad \text{(Ecuación 8)}$$

Donde *n* es el número de ítems en común que presentan los dos usuarios, y el valor de *N* es una constante, que se establecerá en función del domino. En [8] se propone un valor de 50, de forma que si *n* es mayor o igual que *N*, el peso no variará.

Filtrado colaborativo basado en ítem (CF-I)

Para intentar solventar problemas de escalabilidad y dispersión se ha propuesto en [48, 50] una variante de filtrado colaborativo basada en los ítems en vez de en los usuarios, de forma que en vez de estudiarse la similitud entre usuarios y proporcionar predicciones en base a sus votaciones, se estudia el comportamiento de los propios ítems en sí, estableciendo cuáles presentan valoraciones similares y realizando predicciones en base a los propios ítems, y no a los usuarios, creando un modelo de predicción. Por ejemplo, si se ha decidido que los usuarios que votan *Indiana Jones y el templo maldito*, *Indiana Jones y la última cruzada* y *En busca del arca perdida* lo hacen de forma similar, se puede usar las valoraciones que un usuario da para dos de ellas a la hora de estimar la puntuación de la tercera.

La construcción de estos sistemas se realiza de forma análoga a los usuarios sólo que en vez de explorar la matriz de usuarios-ítems por filas para

establecer la similitud entre usuarios, se hace por columnas, para obtener similitudes entre ítems.

Hasta el momento no se ha mencionado, pero los algoritmos de filtrado colaborativo basados en usuario están pensados para trabajar online. Esto quiere decir que se espera que en tiempo real el usuario solicite una recomendación y el sistema realice todos los cálculos necesarios para aportar una recomendación.

Con los algoritmos basados en ítems lo que se pretende es buscar datos que sean computables *offline*, es decir, realizar la mayor parte del cálculo antes de que el usuario solicite una recomendación, de forma que el coste computacional después de la solicitud sea mínimo.

Representación dimensional reducida

La idea que se plantea en [48] para reducir problemas con la dispersión, escalabilidad y con los ítems sinónimos es la de crear una *meta-representación* de los ítems y que las puntuaciones de los usuarios sobre los ítems reales impliquen una puntuación sobre los *meta-ítems*, creados y elegidos en función de la similitud de características entre ítems reales.

Algo similar a esto viene a proponerse en [50], donde se utiliza previamente un filtrado colaborativo basado en ítems para determinar clases de similitud entre ítems y combinarlos posteriormente para determinar la similitud entre usuarios, en lo que ellos consideran una clase de *algoritmos de recomendación de alto orden de interpolación basados en ítems*, definiendo incluso medidas que verifican el grado en el que se consiguen recomendar ítems *ocultos.*

Recomendaciones basadas en reglas de asociación

En [48] se propone una aproximación *basada en reglas* a la hora de realizar la recomendación usando algoritmos de obtención de reglas de asociación para encontrar asociaciones entre ítems relevantes y generar así una recomendación de ítems basada en la fuerza de la asociación entre ítems. Para ello sólo se utilizan los *k* vecinos elegidos, lo cual puede no proporcionar reglas de asociación lo suficientemente fuertes en la práctica, pudiendo generar insuficientes productos a la hora de realizar la recomendación.

Selección de instancias

Para resolver el problema de la escalabilidad, en diversos artículos [27, 49, 51, 59, 60], se proponen distintas aproximaciones de la técnica denominada *selección de instancias*, mediante la cual se eliminan aquellas instancias irrelevantes y/o redundantes del conjunto de entrenamiento.

La motivación de esta extensión proviene de diversos frentes:

- Conforme el número de datos crece se hace más y más costoso para el algoritmo buscar en la base de datos completa. Esto va en perjuicio directo del tiempo de respuesta del sistema.
- ¿Realmente son útiles todas las instancias almacenadas en la base de datos para que el conjunto de entrenamiento considere la suficiente información?
- Si una instancia no está muy bien descrita por sus características, ¿qué clase de impacto tendría dicha instancia a la hora de realizar una predicción?

La idea a la hora de realizar la selección de instancias parte de que existen ciertos ítems para los que se puede encontrar una relación estadística entres sus votaciones, de forma que muchas veces la puntuación de un ítem está directamente relacionada con la puntuación de otro. Por ejemplo, se puede decir que a todos los individuos que han votado de forma positiva las películas *Spiderman*, *Hulk*, *Superman*, y/o *Batman*, estarán predispuestos a votar de forma parecida *Spiderman 2*, existiendo una clara relación entre las votaciones de unos y otros. Igualmente, aquellos que han votado de una forma muy positiva películas como *El apartamento*, *Leyendas de Pasión*, o *Lo que el viento se llevó*, posiblemente voten negativamente *Spiderman 2*.

Si somos capaces de encontrar estas relaciones entre ítems, seremos capaces de encontrar aquellos usuarios que votan un número importante de ítems relacionados con aquél para el que queremos obtener una predicción. Este tipo de instancias serán más relevantes a la hora de realizar estimaciones que aquellas que no realizan este tipo de votaciones. Para encontrar estas relaciones se usa una medida de dependencia estadística utilizada en teoría de la información y denominada *información mutua*.

Se ha demostrado que este método no sólo mejora la exactitud, sino también la eficiencia de los algoritmos de filtrado colaborativo.

Estudio de las características del ítem

En [61, 62] se propone la mejora de la exactitud basándose en el estudio de las características de los ítems, de forma que en vez de considerarse las votaciones para los ítems en sí, se consideran también para sus características.

Tenemos entonces matrices de votaciones en las que para cada usuario aparecen las valoraciones para un ítem, y además para cada una de las características que se han definido para el ítem; el proceso posterior se lleva a cabo teniendo en cuenta todas estas valoraciones.

2.6 Ejemplos de sistemas comerciales basados en filtrado colaborativo

Vamos a ver una serie de ejemplos de sistemas colaborativos que funcionan bastante bien a nivel comercial y que en su momento supusieron un hito en recomendaciones personalizadas.

a) Tapestry

El pionero. Tapestry, un proyecto de Xerox PARC, esta considerado como el primer sistema de recomendación que implementaba filtrado colaborativo. Tapestry permitía a sus usuarios encontrar documentos basados en comentarios hechos previamente por otros usuarios. Al ser un experimento pionero surgieron muchos problemas ya que sólo funcionaba correctamente con pequenos grupos de personas y eran necesarias consultas de palabras específicas para obtener resultados lo que dificultaba en gran medida el propósito último del filtrado colaborativo. También tenía otras carencias, como la falta de privacidad. A pesar de todo fue un sistema que resulto crucial para el posterior fulgurante crecimiento de los sistemas de recomendación colaborativos.

b) Zagat Survey (http://www.zagat.com)

Figura 3. Vista de la interfaz de Zagat Survey

Zagat Survey es una empresa americana fundada en 1979 que se dedica a la edición de todo tipo de guías de restaurantes, hoteles, clubes o tiendas de distintas ciudades de los Estados Unidos y Canadá. En zagat.com los usuarios

registrados pueden votar distintos aspectos (hasta 30) del local referido y, además, introducir pequeños comentarios con su experiencia. En base a estas votaciones los responsables de la empresa asignan sus puntuaciones en sus guías anuales y hacen recomendaciones individuales a sus usuarios a través de su web.

c) Filmaffinity (http://www.filmaffinity.com)

Filmaffinity es un proyecto español de gran proyección internacional que desde 2002 se encarga de recibir puntuaciones de todo tipo de películas y dar recomendaciones a sus usuarios. Su funcionamiento es sencillo: el nuevo usuario puede empezar a votar las películas que haya visto con la ayuda de unos tours dirigidos que atemperan de manera inteligente los problemas de dispersión y del nuevo ítem; posteriormente el sistema calcula las almas gemelas de este nuevo usuario y le recomienda las películas favoritas de estas almas gemelas que no haya votado el usuario. También ofrece la posibilidad de puntuar series de televisión, de ver la información de tus almas gemelas y contactar con ellas, realizar críticas y comentarios sobre las películas y series vistas y consultar los mejores ítems por géneros, décadas o en general.

Figura 4. Página principal de Filmaffinity

d) Last.fm (http://www.last.fm)

La popularidad adquirida en los últimos tiempos por los sistemas de recomendación colaborativos ha propiciado que se haya ampliado el rango de aplicación de los mismos. Una nueva generación de sistemas de recomendación colaborativos ha surgido en el ámbito de la radio musical vía Internet.

Last.fm es un sistema de recomendación musical además de una red social y una radio a través de Internet. Cada nuevo usuario va creando su propio

perfil de manera muy sencilla: el usuario puede escuchar las radios personalizadas (canciones favoritas) del resto de usuarios y decidir si les gustan (pasan a formar parte del perfil del propio usuario) o si por el contrario no quiere volverlas a escuchar. Con este perfil y gracias a un algoritmo de filtrado colaborativo el sistema va cerrando el cerco sobre los usuarios con gustos mas afines con el usuario activo (los denominados vecinos) y recomendando grupos y artistas que no se encuentran en su perfil pero que si son favoritos dentro de su vecinos. Estas acciones son continuas por lo que la calidad de las recomendaciones se refina de manera ciertamente importante cuando se es un usuario veterano dentro la aplicación.

Figura 5. Página de recomendaciones de Last.fm

Sin duda, Last.fm y otras aplicaciones de similares características, están siendo una de las grandes revoluciones de los últimos tiempos en Internet gracias a su marcado carácter social, a su amplio catalogo musical y a la calidad de sus recomendaciones.

e) MovieLens (http://movielens.umn.edu)

MovieLens es un sistema de recomendación de películas online basado en filtrado colaborativo. Desarrollado por el GroupLens Research de la Universidad de Minnesota (http://www.grouplens.org), recolecta puntuaciones sobre películas de sus usuarios y en base a esos datos agrupa los usuarios de similares gustos. Atendiendo a las puntuaciones de todos los usuarios dentro de un grupo se intenta predecir para cada usuario individual su opinión sobre películas que todavía no ha visto.

Figura 6. Interfaz de MovieLens

En un principio utilizaba algoritmos basados en usuario para realizar sus predicciones y recomendaciones pero desde hace un tiempo emplea algoritmos basados en ítem porque le ofrecen mejores resultados.

Los datos sobre sus usuarios y sus puntaciones son privados pero los investigadores de GroupLens mantienen como publicas dos ejemplos de 100000 y un millón de puntaciones respectivamente. Estos ejemplos se pueden descargar desde la propia pagina web (http://www.grouplens.org) para realizar pruebas con diversos algoritmos colaborativos.

f) Where to Cycle (http://www.wheretocycle.com/)

Esta comunidad de usuarios se dedica a compartir información sobre viajes realizados en bicicleta por todo el mundo. Se dedica a fines no comerciales e intentan localizar los mejores lugares para viajar en bicicleta. Permite realizar valoraciones sobre los lugares visitados por el usuario así como realizar las recomendaciones más adecuadas mediante filtrado colaborativo.

Figura 7. Página de recomendaciones de wheretocycle.com

g) 2DoNext (http://www.2donext.com/)

Figura 8. Página principal de 2donext.com

Después de años de trabajo, Mike se decidió a colgar su sitio web. El sitio web se llama 2DoNext, la idea de este sitio es permitir a la gente puntuar películas, videojuegos, libros, cómics y música por separado. También ofrece la posibilidad de compartir experiencias con familiares, amigos y todo tipo de personas que han puntuados ítems similares.

El sitio también incluye un sistema basado en filtrado colaborativo que recomendará de acuerdo con lo que el propio usuario ha estado puntuando. El objetivo es permitir conectarse con otros usuarios y puntuar sus aficiones, de tal forma que se les pueda recomendar acerca de que es lo próximo que podrían hacer para entretenerse.

h) Pure Video (http://purevideo.com/)

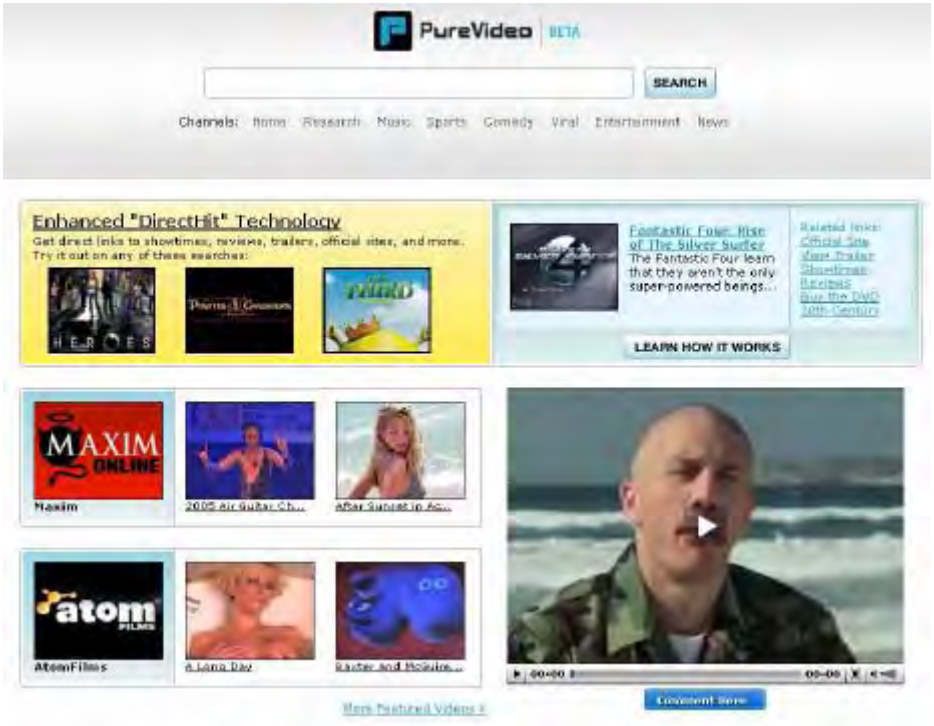


Figura 9. Página principal de purevideo.com

PureVideo Networks, Inc. es una compañía dedicada a las nuevas tecnologías que permite descubrir y promocionar videos online. Fue fundada por Softbank Capital, en Mayo de 2005, como consecuencia de la revolución de los videos online generados por los propios usuarios.

Con los videos online como un medio al alza en el mundo de la web 2.0, son muchos los sitios web de videos en el Mercado. El motor de búsqueda PureVideo fue recientemente puesto en marcha para intentar ayudar y organizar

los espacios de videos online. En este caso el filtrado colaborativo y el perfil de usuario deberían avanzar bastante de manera que el usuario sea capaz de acceder a su lista de archivos multimedia desde cualquier sitio, o buscar nuevos archivos de una manera fácil.

i) Findory (http://findory.com/)

Findory es un sitio de noticias de actualidad que emplea filtrado colaborativo para poder mostrarle al usuario noticias de su interés. Fue desarrollado por Greg Linden en Seattle, proveniente de Amazon (el rey de los sitios de recomendaciones).

Figura 10. Página de inicio de Findory

j) Silver Egg (http://www.silveregg.co.jp/en/)

Silver Egg proporciona servicios web comerciales que funcionan con tecnologías basadas en Inteligencia Artificial. El servicio más importante es Aigent, un servicio en ASP empleado en las páginas web lideres de comercio electrónico para realizar recomendaciones personalizadas sobre productos. Utiliza un modelo de filtrado colaborativo basado en redes bayesianas para elegir los productos que se van a recomendar, adaptando continuamente las recomendaciones a los cambios de compra del usuario.

Figura 11. Página de inicio de Silver Egg

k) Amazon (http://www.amazon.com)

Figura 12. Página personalizada en Amazon

Esta página web contiene un sistema de recomendación que mezcla técnicas colaborativas con aquellas basadas en contenido, guardando las preferencias del usuario activo y combinándolas con objetos relevantes para genera recomendaciones (las ya célebres del tipo "*La gente que compró este producto también compró estos…*").

Otros sistemas de recomendacion colaborativos (tanto de ámbito comercial como no comercial) que pueden ser de interés para que el lector compruebe el funcionamiento y utilidad de los mismos son los siguientes: Pandora (http://www.pandora.com), Live365 (http://www.live365.com), TiVo (http://www.tivo.com), aNobii (http://www.anobii.com/anobi), y otros muchos, como los que pueden ser consultados en la siguiente dirección (http://en.wikipedia.org/wiki/Collaborative_filtering).

3. ANÁLISIS DEL FC EN LA RECOMENDACIÓN DE ITINERARIOS ACADÉMICOS Y ASIGNATURAS PARA EL BACHILLERATO

3 ANÁLISIS DEL FC EN LA RECOMENDACIÓN DE ITINERARIOS ACADÉMICOS Y ASIGNATURAS PARA EL BACHILLERATO

Tras hacer un repaso a las posibilidades que nos ofrece el campo de los algoritmos de filtrado colaborativo, vamos a proceder al estudio del dominio concreto en el que los pretendemos aplicar, realizando experimentos para aquellas de las alternativas presentadas que prometen mejores resultados dadas las características propias de los datos e ítems observados.

Como hemos visto anteriormente, el propósito principal de esta memoria de investigación es el de evaluar el funcionamiento de los sistemas de recomendación basados en filtrado colaborativo a la hora proporcionar recomendaciones al alumnado sobre qué asignaturas serán las más adecuadas en función de su expediente académico.

Más concretamente, el sistema deberá recomendar una serie de asignaturas optativas al alumno usuario, que serán consideradas como las asignaturas que el alumno debería elegir para el curso que va a comenzar. Como en todo sistema de recomendación, el alumno tendrá libertad de elegir o no esas asignaturas, pero la herramienta serviría de apoyo a la hora de tomar tal decisión. El sistema orientaría al alumnado en base a los datos obtenidos.

Sin embargo, es interesante también observar la exactitud de las predicciones no sólo para asignaturas optativas, sino también para el resto, de forma que en un futuro pueda plantearse un sistema más general, y analizar para qué otras tareas podrían utilizarse las predicciones, como por ejemplo la recomendación de perfiles académicos.

Pasemos a ver una breve descripción del ámbito educativo en el que nos moveremos.

3.1 Introducción al contexto del Bachillerato

Aunque esta información está más completa en el *Anexo II*, vamos a hacer una breve introducción sobre el Bachillerato para poder situarnos en el contexto del problema.

El Bachillerato es la última etapa de la Educación Secundaria, es de carácter post-obligatorio (voluntario) y su duración es de dos cursos, normalmente entre los 16 y los 18 años.

Tiene modalidades diferentes que permiten una preparación especializada de los alumnos (con elección de distintos itinerarios dentro de cada modalidad) para su incorporación a estudios superiores o a la vida activa.

La forma normal de acceso al Bachillerato pasa por estar en posesión del título de Graduado en Educación Secundaria o título equivalente, y para ello es necesario superar 4º de E.S.O.

3.1.1 Modalidades e itinerarios del Bachillerato

Las enseñanzas de Bachillerato están estructuradas en modalidades; unas de corte más académico y otras más profesional, facilitando así que cada alumno pueda elegir su propio itinerario formativo en función de sus capacidades e intereses académicos y profesionales.

El Bachillerato actual comprende cuatro modalidades, que son las siguientes:

a. Artes
b. Ciencias de la Naturaleza y de la Salud
c. Humanidades y Ciencias Sociales
d. Tecnología

El Bachillerato permite a los alumnos cursar estos estudios de acuerdo con sus preferencias, en virtud de la elección de una modalidad entre las cuatro previstas, una opción dentro su modalidad y unas determinadas materias optativas. Estas sucesivas elecciones configuran el itinerario personal de cada alumno. Aquí es donde nosotros vamos a evaluar hasta qué punto puede un sistema de recomendación colaborativo orientar al alumnado a la hora de conformar su itinerario personalizado.

La posibilidad de elección aumenta progresivamente de primero a segundo curso, en el que se incluyen, dentro de cada modalidad, 2 ó 3 opciones posibles que guardan, en la mayor parte de los casos, una estrecha relación con determinados estudios posteriores universitarios o profesionales.

La elección de determinadas materias optativas ofrece, por otra parte, bien la posibilidad, a los alumnos que lo desean, de concurrir a las pruebas de acceso a la universidad por más de una opción, o por una opción distinta de la prevista para la modalidad o itinerario cursado, bien la posibilidad de profundizar en aspectos específicos de su modalidad, o bien ampliar la formación que pueden adquirir en la modalidad elegida, cursando como optativas materias propias de otras modalidades o materias optativas de posible interés formativo para los alumnos de todas las modalidades del Bachillerato.

3.1.2 Organización de las Materias

El Bachillerato se organiza en:

1. materias **comunes**, para todos los alumnos independientemente de la modalidad elegida. Pretenden contribuir a la formación general del alumnado

2. materias **propias** de cada modalidad, que caracterizan a cada una de las modalidades y contribuyen a que el alumno obtenga una formación específica ligada a la modalidad elegida, y
3. materias **optativas**, que amplían la posibilidad de elección.

Los alumnos deberán cursar seis materias propias de la modalidad elegida, tres en cada curso.

Las distintas Administraciones Autonómicas y los propios centros educativos tienen ciertas libertades a la hora de organizar tanto las Modalidades como las materias optativas ofertadas, que pueden variar de un centro educativo a otro.

Además, los alumnos podrán elegir como materias optativas no sólo las que resulten de lo previsto en el apartado anterior, sino también cualesquiera de las materias definidas como propias de las diferentes Modalidades, de acuerdo con lo que al efecto determinen las Administraciones educativas en función de las posibilidades de organización de los Centros.

Como vemos, al alumnado se le requiere el tomar una serie de decisiones que, por experiencia propia y observación en los centros en los que he trabajado, generalmente no está preparado para tomar o se encuentra con numerosas dudas, y en muchos casos tales decisiones se realizan teniendo en cuenta más el grado de dificultad de las materias, de manera que se suele optar en algunos casos por aquellas asignaturas más fáciles en vez de por las que más convienen para el futuro académico o profesional.

En concreto, el alumno debe realizar 3 tipos de decisiones:

1. escoger la modalidad
2. escoger que 3 asignaturas de la modalidad estudiar cada curso
3. escoger un número variable de optativas en cada curso de entre las ofertadas por el centro, incluyendo como posibilidad el cursar como optativas otras materias propias de la modalidad

Para ayudar al alumnado a tomar estas decisiones sería ideal disponer una herramienta que ayudara a:

- Explorar el conjunto de modalidades y sus proyecciones
- Estudiar las distintas alternativas una vez elegida la modalidad
- Estudiar las posibilidades disponibles a la hora de escoger materias optativas
- Proponer uno o varios itinerarios educativos

De aquí es de donde surge la idea y motivación de nuestro estudio, por lo que vamos a pasar a delimitar nuestras intenciones y propósitos de cara al resto del trabajo.

3.2 Objetivo del estudio

En nuestro estudio pretendemos analizar las posibilidades de predicción que presentaría un hipotético sistema de recomendación basado en filtrado colaborativo que trabajaría con la información relativa a los expedientes académicos del alumnado. El sistema colaborativo trataría de comparar expedientes entre alumnos y de predecir qué calificaciones obtendría un alumno concreto en función de las obtenidas hasta el momento y en función de las calificaciones que obtuvieron otros alumnos en el pasado con un expediente similar.

La pretensión inicial es la de comprobar hasta qué punto tal sistema podría orientar al alumnado a la hora de recomendar itinerarios educativos o asignaturas optativas concretas.

Es importante hacer notar que queremos probar este tipo de técnicas a la hora de orientar al alumnado porque asumimos que las calificaciones que se van teniendo a lo largo de la formación académica de cada individuo tienen bastante que ver con diversos factores como el nivel de conocimientos, las aptitudes y las preferencias del alumnado. Puesto que nuestro objetivo es tan simple como comprobar si el uso de un sistema de filtrado colaborativo es beneficioso en este ámbito, no haremos inicialmente otros estudios preliminares (minería de datos, estudios cualitativos y/o demográficos, etc.).

Para nuestro estudio vamos a utilizar los algoritmos colaborativos más extendidos y se tratarán de ajustar al dominio, de optimizar e incluso se estudiará la posible inclusión de mejoras propias.

Una vez obtenidos los resultados serán analizados convenientemente y, si las conclusiones son positivas y las predicciones prometedoras, se construirá un sistema que permita realizar estas recomendaciones.

Según la formulación de los sistemas de filtrado colaborativo, para conseguir que un algoritmo de este tipo funcione debemos [9]:

- Determinar la tarea principal del sistema.
- Analizar los datos con el que vamos a trabajar, las características del dominio, sus propiedades y peculiaridades.
- Profundizar en el espacio del problema, decidiendo los datos a utilizar, su proveniencia y las características del dominio.
- Evaluar las distintas posibilidades de elección de vecinos y de realización de predicciones. Para ello también necesitaremos:

  - seleccionar un conjunto de datos para la evaluación,
  - elaborar un conjunto de pruebas y aplicar métricas de calidad,
  - estudiar los resultados y comprobar la viabilidad del sistema.
- Por último y después de decidir si el sistema es o no viable, descubrir posibles implicaciones, aplicaciones y alcanzar aquellas conclusiones que se deriven de todo el proceso.

3.3 Tarea principal

El objetivo que perseguimos en este estudio es diseñar un algoritmo que nos permita *encontrar todos los ítems buenos*, en concreto, puesto que tratamos de encontrar aquellas materias para las que se prevé que el alumno obtendría resultados exitosos, se adapten a sus conocimientos y/o aptitudes, etc.

Tengamos en cuenta que esto no sería la meta final, sino un paso previo a la hora de realizar la recomendación puesto que, una vez estimado este conjunto de materias, habría que analizar el modo en el que los resultados se le mostrarían al alumno para orientarlo de la mejor forma posible y sin perder de vista que siempre sería él quien tendría la última palabra a la hora de escoger.

Una vez aclarado esto, se hace necesario indicar que el dominio concreto en el que nos movemos es bastante particular y específico. Esto a priori puede interpretarse como una fuente de problemas, pero veremos que en realidad estas características y peculiaridades en la mayoría de los casos nos facilitan nuestra tarea, e incluso evitan algunos de los problemas propios del CF explicados con anterioridad.

3.4 Análisis de los datos

Como tarea básica a realizar, debemos estimar la calificación que, en base a la experiencia de otros alumnos, el alumno activo obtendría en una serie de asignaturas en caso de cursarlas.

Las valoraciones en nuestro dominio son las calificaciones que obtiene el alumnado en las distintas asignaturas que cursa, con lo que las ternas *usuario-ítem-votación* en nuestro dominio pasan a tomar la forma de ternas del tipo *alumno-materia-calificación.*

Algo a tener muy en cuenta, porque elimina posibles fuentes de problemas, es que el espacio formado por los pares *alumno-materia* resulta drásticamente menor que en la mayoría de los sistemas comerciales existentes. Aunque se puede almacenar información relativa a gran cantidad de alumnos, lo que hace singular a nuestro problema es el hecho de que la cantidad de asignaturas (ítems) a contemplar es, para el caso de un alumno concreto, considerablemente pequeña.

Pongámonos en el caso de un alumno de 4º de ESO que quiere decidir qué asignaturas escoger en 1º de Bachillerato. Contemplando un máximo estimado de 20 asignaturas por curso, contando las optativas, el número total de asignaturas que se barajarían sería de 80 materias si tenemos en cuenta desde 1º de ESO hasta 4º de ESO, por ejemplo, y con esto bastaría para establecer el grado de similitud entre alumnos.

Y continuando con el proceso, para realizar una predicción no se necesitarían predecir todas las asignaturas que el alumno no hubiera cursado, sino sólo aquellas que pertenecieran a 1º de Bachillerato, el curso al que pretende acceder, es decir, unas 30 o 40,

Estas cuestiones facilitan enormemente el trabajo a priori, y nos asegura que los problemas de *escalabilidad* serán mínimos.

3.4.1 Características del dominio

Puesto que la tarea o tareas que se pretenden resolver han quedado definidas claramente, vamos a pasar a estudiar el resto de cuestiones definidas en el punto *2.2.1*:

Trasfondo y contexto del ítem a recomendar

Los ítems con los que vamos a trabajar son materias, asignaturas; concretamente vamos a centrarnos en aquellas pertenecientes al Bachillerato del sistema educativo actual. La razón de ceñirnos únicamente a estas materias viene determinada por diversas cuestiones:

- Ha resultado difícil obtener datos reales que permitan realizar experimentos de confianza para cursos anteriores; los únicos cuyo volumen ofrecen suficientes garantías han sido los de 1º y 2º de Bachillerato, y en menor medida de 4º de E.S.O.
- Estos cursos son mucho más interesantes que otros anteriores a la hora de determinar el grado de éxito en el que los algoritmos de CF puedan utilizarse en tal dominio, puesto que las materias que se imparten ofrecen una mayor variedad tanto en temática como en tipología, y son más numerosas que en el resto de los cursos .

Se ha considerado imprescindible tener en cuenta ciertas características específicas de las materias en el Bachillerato anteriormente mencionadas, para establecer agrupaciones de asignaturas:

1. La **tipología**: se han agrupado por tipología para poder estudiar los resultados no sólo de forma global, sino también de forma más específica, pudiendo distinguir entre las materias comunes, las propias de modalidad y las optativas.
2. El **curso**: cada asignatura deberá ser contemplada dentro del curso en el que se imparte.

3. La **modalidad**: las materias de modalidad contemplan a qué modalidad se refieren (*Artes*, *Ciencias de la Naturaleza y de la Salud*, *Humanidades y Ciencias Sociales* y *Tecnología*). Pueden existir asignaturas compartidas por varias modalidades.

Profundizando en la tipología que hemos mencionado, en Bachillerato existen 3 tipos de asignaturas:

- **Comunes**: son obligatorias y todas las modalidades las contemplan
- **Propias** o de modalidad: específicas de la modalidad que se está cursando
- **Optativas**: de libre elección por parte del alumnado, independientemente de la modalidad

Con respecto al tratamiento que se debe dar al alumnado, sólo es necesario tener en cuenta las calificaciones correspondientes a las asignaturas cursadas. Con los datos que de ellas se derivan se puede hacer una predicción consecuente, simplemente proporcionando el curso para el que se quiere realizar la predicción.

Novedad frente a calidad en la recomendación

En nuestro caso concreto, es completamente indispensable primar la calidad con respecto a la novedad. Aunque, desde un punto de vista realista, las materias susceptibles de ser recomendadas para un alumno son siempre nuevas para él, a no ser que haya repetido curso, por lo que al hablar de novedad habría que matizar y decir que serían aquellas que el alumno no se había planteado cursar en ningún momento.

También es verdad que en algunos de los ámbitos específicos en los que se podría aplicar el algoritmo que nos atañe, quizá si tuviera cierta importancia la novedad, entendida como hemos matizado en el párrafo anterior. Por ejemplo, si nos encontráramos en el ámbito universitario, puede darse el caso de que haya asignaturas poco conocidas por el alumnado, o bien desconocidas en el sentido de que no se sepa bien su contenido o si uno debería o no matricularse de ellas; en este caso, un sistema según podríamos plantearlo quizá resultaría provechoso aportando novedad de modo tal que el alumno pudiera contemplar como posibilidad a la hora de matricularse asignaturas que de cualquier otra manera no se habría planteado.

De cualquier manera, lo que si debe quedar claro es que la calidad en las recomendaciones es indispensable debido al ámbito en el que nos movemos y a la importancia de las decisiones a tomar.

Análisis del coste/beneficio con respecto a los falsos/verdaderos positivos/negativos

Como acabamos de ver, es muy importante la calidad de las recomendaciones, sobre todo debido a la relación coste/beneficio que puedan tener los falsos/verdaderos positivos/negativos. Vamos a explicar cuáles serían nuestros falsos positivos/negativos y su coste, que son los que más problemas pueden acarrearnos:

- **Falso positivo**: un falso positivo sería la recomendación de una asignatura considerada como adecuada, y en la que realmente el alumno no obtiene buenos resultados, o bien no sirve para sus futuras intenciones. Es decir, una recomendación en la que el sistema falla. También es de gran importancia el intentar evitar en la medida de lo posible los falsos positivos, puesto que pueden desembocar en el peder la oportunidad de cursar materias más beneficiosas para el alumno.
- **Falso negativo**: estaríamos en el caso de asignaturas que se recomendaría no escoger, o que requerirían actividades de refuerzo. Este caso es el menos perjudicial de los dos falsos, puesto que nunca está mal el realizar actividades de refuerzo (siempre y cuando no perjudiquen otras actividades), aunque en el caso de que se hubiera dado como mala una asignatura una que resultara beneficiosa para el alumno, el problema sería algo mayor, y de difícil vuelta atrás.

Granularidad de las valoraciones del usuario

Como vimos, la granularidad se corresponde aquí a las preferencias reales del alumno. En el problema que estamos tratando, estas preferencias reales no son calificaciones, sino el grado de satisfacción con el que se concluiría el estudio de la asignatura en cuestión, por lo que debemos realizar ciertas consideraciones.

Podemos considerar que un alumno estará satisfecho con la recomendación del sistema si se da uno de los siguientes casos:

a) la asignatura le ha gustado,

b) la asignatura era conveniente para su formación,

c) ha sacado una puntuación alta en la asignatura recomendada,

d) ha obtenido una puntuación positiva en una materia en la que se recomendó refuerzo

En los casos contrarios a los anteriores, es de suponer que el alumno no estará de acuerdo con la recomendación.

Podría entonces medirse el grado de satisfacción del alumnado en una escala, por ejemplo, de 1 a 5, con etiquetas lingüísticas que expresaran el grado de satisfacción del alumno.

Sin embargo, nosotros sólo disponemos actualmente de las propias calificaciones del alumnado, por lo que a la hora de considerar el grado de satisfacción del alumnado deberemos ceñirnos a esas calificaciones, de 0 a 10, donde los valores más altos podrían significar un alto grado de satisfacción, mientras los más bajos, alto grado de disgusto con respecto a la asignatura.

3.4.2 Características inherentes

Hoy día, en cualquier centro de cualquier nivel al que queramos referirnos, los datos relativos a alumnos, materias y calificaciones suelen estar almacenados de forma informatizada, por lo que el camino que debemos seguir queda bastante definido al tener ciertas propiedades prefijadas.

Valoraciones implícitas, explícitas, o ambas.

Los datos de las valoraciones con las que el sistema trabajará no pueden considerarse ni estrictamente implícitos, ni tampoco explícitos. El alumnado no aportará los datos directamente, sino que estos serán objetivos y estarán directamente proporcionados por el personal encargado de calificar al alumno, derivados de su actuación frente a la materia y en base a los criterios definidos en la programación didáctica de la asignatura.

Desde cierto punto de vista podría considerarse que los datos son explícitos, dado que en realidad alguien debe introducirlos en el sistema de forma manual, y no son obtenidos automáticamente por éste; sin embargo, y desde mi punto de vista, para poder considerar unos datos explícitos el usuario debería tener algún mecanismo para poder modificarlos directamente, es decir, el usuario debería poder expresar sus propias percepciones mediante las valoraciones, mientras que en este caso las calificaciones dependen del juicio de terceras personas. De esta forma, y aunque es cierto que dependen directamente tanto del comportamiento como de los conocimientos del alumno con respecto a la asignatura, es el profesorado quien tiene la última palabra a la hora de calificar. Es más, difícilmente serán modificadas las calificaciones de los alumnos a posteriori, hecho que presenta ciertas ventajas al enfrentarnos a un sistema poco variable con el tiempo.

Por todo esto debemos darnos cuenta de la peculiaridad de los datos con los que estamos tratando puesto que, estrictamente hablando, no podríamos considerarlos implícitos teniendo en cuenta el matiz que se suele dar a este tipo de datos en otros sistemas, en los que los datos implícitos son aquellos derivados del comportamiento del usuario y recogidos de forma automática en base al análisis que el sistema hace de ese comportamiento. En nuestro caso, el sistema

no realiza ningún análisis, le viene ya hecho. Debemos pues valorar el impacto que los sistemas basados en CF pueden tener en datos tan particulares.

Se hace entonces muy importante también comprobar el impacto de los algoritmos de filtrado colaborativo en unos datos tan peculiares y en gran medida distintos a lo que se ha venido trabajando hasta el momento en el ámbito de los sistemas colaborativos.

En otro orden de cosas, hay que hacer notar que las valoraciones del usuario son *obligadas*. Esto quiere decir que el alumno no vota aquellas materias que el estime conveniente, ni de la forma que crea mejor, sino que inevitablemente se produce una valoración, que además es externa al alumno, para cada una de las materias que ha cursado. De esta forma no existirá ningún alumno que tenga un número inferior de asignaturas puntuadas que aquellas que se cursen como mínimo en un año. Gracias a esto, difícilmente nos encontraremos con problemas de *dispersión* o de aquellos que presentan los *usuarios nuevos*.

Escala y dimensión de las valoraciones.

Las asignaturas ya presentan una escala de valoraciones nominadas predefinida. Esta escala además tiene una correspondencia numérica que es la que vamos a utilizar en el conjunto de datos:

- Insuficiente: de 0 a 4 ambos inclusive
- Suficiente: 5
- Bien: 6
- Notable: 7 y 8
- Sobresaliente: 9 y 10

Las valoraciones serán entonces numéricas y enteras, tomando valores desde el 0 hasta el 10, ambos inclusive.

Sólo se tomará en cuenta la calificación objetiva de una asignatura, por lo que las valoraciones serán unidimensionales.

Presencia de marca temporal.

Este es un punto interesante a tratar, puesto que existe la posibilidad de que algunos alumnos tengan varias calificaciones para la misma asignatura. Esto se explica si un alumno suspende una materia: posteriormente tiene que recuperarla. El conjunto de datos tendría entonces como marca temporal el curso y la convocatoria en la que se produce cada calificación del alumno sobre la asignatura. Más adelante veremos las implicaciones que esta *marca temporal* puede conllevar.

3.4.3 Resumen del análisis de las características del dominio

Como conclusión de este análisis previo de los datos, podemos decir lo siguiente:

- Tendremos diversas materias, cada una de las cuales tendrá un tipo (optativa, propia o común), un curso, y una modalidad en el caso de las optativas.
- No se primará la novedad que puedan aportar las asignaturas recomendadas, haciendo el mayor hincapié posible en la calidad de las recomendaciones.
- Las decisiones a las que el sistema intenta ayudar son suficientemente relevantes como para no permitir tolerar falsos positivos, y minimizar lo máximo posible los falsos negativos.
- Las valoraciones se recogerán de forma implícita de los datos relativos al expediente del alumno.
- Las calificaciones tomarán su valor de una escala formada por números enteros que irán del 0 al 10.
- Una asignatura puede ser calificada en distintos momentos, concretamente en distintos cursos académicos, para un alumno concreto, mientras que dicho alumno no supere la asignatura.
- Existirán escasos problemas de *dispersión*, puesto que el alumno está liberado de responsabilidad a la hora de aportar datos, y la evaluación de las asignaturas es obligada, lo que garantiza que tendremos un número mínimo de valoraciones.
- Los datos son objetivos, proporcionados por el profesor encargado, por lo que nos aseguramos el no encontrarnos con problemas de *robustez* [63].
- Por otro lado, el número de materias a manejar es relativamente pequeño, algo muy deseable en este tipo de sistemas, eludiendo de forma considerable problemas relativos a la *escalabilidad* de nuestro sistema.
- Para terminar, no se dará el caso de que se espere una recomendación para un alumno con pocas o ninguna calificación, lo que hace que el cálculo de similitudes entre alumnos sea más viable y de mayor confianza que en otros ámbitos.
- Con respecto al problema del nuevo-usuario, de ceñirnos a la predicción de calificaciones en materias optativas, deberíamos decir que en la práctica es imposible que se de esta situación, puesto que el primer curso en el que se pueden elegir optativas es en 3º de ESO, por

lo que el alumno de forma obligatoria ha cursado 1º y 2º de ESO con anterioridad, de modo que como mínimo tendremos información relativa a las materias correspondientes a esos cursos.

- Con respecto al problema del nuevo-ítem, sin embargo, debemos hacer notar que aunque es muy difícil el que aparezcan nuevas asignaturas, dado que la creación de éstas viene determinada oficial y legalmente, el sistema, para poder realizar predicciones de una nueva asignatura, tendría que esperar a que un número suficiente de alumnos la cursaran, o bien habría que contemplar la posibilidad de ampliar el sistema con algún mecanismo demográfico, basado en contenido o en conocimiento.

3.5 Particularidades del dominio

A parte de las características propias del dominio que acabamos de ver, se hace necesario explicar a parte una serie de particularidades que este complejo dominio presenta a la hora de realizar las predicciones.

En primer lugar, y como ya se ha hecho notar anteriormente, existe la posibilidad de que haya individuos con **varias calificaciones** para una misma materia, debido a que la haya suspendido y haya tenido que presentarse en años sucesivos. Esto presenta dos problemas: uno a la hora de establecer la similitud entre vecinos, y otro a la hora de realizar predicciones.

A la hora de calcular la similitud entre vecinos, si tenemos varias valoraciones para un mismo usuario, es posible que se evalúe un objeto dos o más veces con respecto a los mismos usuarios, alterando los grados de similitud entre alumnos.

De estas valoraciones sabemos que cumplen las siguientes afirmaciones:

- Sólo existirá una única calificación con valor igual o mayor que 5 por alumno y asignatura.
- En caso de haber varias calificaciones, éstas deben estar restringidas a cursos distintos.
- En caso de haber varias calificaciones, la última de ellas será mayor o igual que 5; todas las anteriores deben por fuerza ser menores o iguales a 4.

Sabiendo esto, existe la posibilidad o bien de contemplar todas las calificaciones (puesto que forman parte del expediente académico del alumno, y por tanto de su historia), o únicamente tener en cuenta la última.

Con respecto las calificaciones de las asignaturas nos surge otro problema propio del dominio: existirán calificaciones digamos normales, con valores entre 0 y 10, pero también existe la posibilidad de que tomen un valor peculiar que se sale de este rango: NP o No presentado. Este valor se produce cuando el profesor

no ha tenido ninguna posibilidad de evaluar al alumno, o bien porque este no ha asistido a clase, o bien por no haber realizado los exámenes o actividades propuestas.

El alumno está matriculado en la materia, y sin embargo no tiene calificación. Un valor de NP generalmente está correspondido con un valor de Suspenso, más concretamente a la hora de introducir valores numéricos suele transcribirse como 0 ó 1, ya que generalmente lo que el alumno hace es no asistir directamente a la asignatura o a las evaluaciones.

En cualquier caso, cualquiera de los dos valores expresa en la práctica un valor de Suspenso, por lo que se nos presenta una disyuntiva: ¿deberíamos asignar realmente tal valor a las calificaciones con NP, o deberíamos considerar las asignaturas en las que se producen como que no presentan calificación?

Con respecto al cálculo de vecinos, si no consideramos que sean calificaciones válidas (sin calificación), el sistema se comportaría como si fueran *productos no evaluados*, por lo que serían ítems que no se tendrían en cuenta a la hora de calcular las similitudes entre usuarios. La otra alternativa sería considerar la asignación de un valor por defecto para estas calificaciones, que expresara suspenso.

Otro posible problema se nos presentaría a la hora del cálculo de predicciones: si no tiene valor la asignatura para un vecino, ese vecino no aporta información. En cambio, si asignamos un valor a la calificación de esa asignatura, tal valor tomaría efecto real en la predicción.

Como decisión, desde mi punto de vista es más práctica la de asignar una calificación por defecto en tales asignaturas, por las siguientes razones:

- Las calificaciones consideradas como NP en casi la totalidad de los casos implica que el alumno no se ha presentado posiblemente porque creía no iba a superar la asignatura, es decir, habría suspendido. Existen excepciones (enfermedades prolongadas, viajes ineludibles, etc.) pero siempre subjetivas y que no cambian el resultado de suspenso.
- El hecho de asignar un valor de suspenso a estas materias, cuando en el 90-95% de los casos realmente serían suspensos, aporta una información bastante relevante a la hora de calcular vecinos, y por supuesto, a la hora de calcular predicciones. Si estuviéramos hablando de alumnos universitarios, un NP es equivalente a un suspenso y la asignatura no aparecería en el expediente como superada.
- En cualquiera de los casos, sea cual sea la razón, el alumno debería repetir la asignatura, o simplemente no se la daría por aprobada, por lo que el efecto práctico es el de un suspenso.

Estas tres razones se consideran suficientes para proponer la asignación de un valor de suspenso a este tipo de calificaciones, y como valor se asignará un 0, El motivo de esto es doble: primero, porque el no asistir a la asignatura o no presentarse a las evaluaciones implica el conocimiento escaso o nulo del individuo en la materia; segundo, existe un desinterés inherente al hecho de no presentarse a una asignatura, que puede expresar disgusto con la misma, por lo que en este caso podría considerarse que el 0 está midiendo el grado de desaprobación que el alumno presenta hacia la asignatura.

Posteriormente y una vez obtenidos los resultados de los experimentos se podrá discutir sobre la adecuación o no de estas medidas.

Como dato técnico a tener en cuenta, y tras observar y clasificar las distintas materias que existen para 1º y 2º de Bachillerato, se ha comprobado que en el conjunto de datos a utilizar en la evaluación del algoritmo, existen materias para los mismos cursos que pese a ser las mismas y denominarse igual, se tratan de forma diferente en función de la modalidad a la que pertenecen. Por ejemplo, la asignatura Inglés Segundo Idioma de 1º de Bachillerato de Artes, se almacena en los datos como una instancia a parte de la asignatura Inglés Segundo Idioma de 1º de Bachillerato de Ciencias de la Naturaleza y de la Salud, cuando en la práctica esas asignaturas son tratadas de forma exactamente igual.

Para evitar los efectos negativos que estos *ítems sinónimos* pueden producir a la hora del cálculo de vecinos y de la generación de predicciones se va a reconvertir el espacio de ítems en otro en el que aquellas materias que hagan referencia en la práctica a la misma asignatura serán consideradas como una única materia a efectos prácticos.

Por último, un hecho curioso y que se prevé va a dar ciertos problemas es el de aquellos alumnos que durante el transcurso del año abandonan el curso académico. Estos son individuos que se matriculan y durante el año pierden el interés, o bien alcanzan edad para proponerse otras vías de estudio (ciclos formativos) o trabajo, y deciden anular la matrícula. El resultado de este alumnado en las calificaciones es de NP o de suspenso. Esto puede acarrear consecuencias bastante perniciosas para el sistema.

En la Tabla 3 tenemos un ejemplo clarificador. El alumno Juan asiste a 1º de Bachillerato regularmente con calificaciones aceptables, sin embargo en 2º decide que no le gusta estudiar y abandona el curso. Al año siguiente aparece Inés, que quiere saber qué asignaturas de 2º de Bachillerato debería coger. Tras realizar una selección de vecinos vemos que los que han sacado unas calificaciones más parecidas a Inés son Juan y Ana, por lo que la predicción se basaría en ellos. Es intuitivo el observar que tal estimación no va a ser muy

acertada y que el hecho de que Juan abandonara el curso va a influir en las predicciones bajando las estimaciones sustancialmente.

	1º de Bachillerato			2º de Bachillerato	
	Informática	Dibujo Artístico I	Dibujo Técnico I	Dibujo Artístico II	Dibujo Técnico II
Ana	9	10	9	9	8
Eduardo	4	8	6	8	4
María	7	4	5	5	6
Juan	7	7	8	1	0
Inés	9	8	8	?	?

TABLA 3: Ejemplo del efecto que producen a la hora de las predicciones los alumnos que abandonan el curso.

Frente a esta clase de alumnado que abandona el curso sin terminarlo se nos presenta la ventaja de que se detectan de forma fácil. Independientemente de los distintos perfiles que puedan tener, suelen presentar una característica común: todos, y para el último año que han cursado (no necesariamente para el resto) presentan una calificación media muy baja con respecto al resto. Generalmente, cuando un alumno abandona un curso presenta en el curso que se produce el abandono unas calificaciones que suelen tomar valores de 0, 1, 2, o como mucho 3 (sin contar aquellas que se consideran como NP).

Podríamos basarnos sólo en la calificación media a la hora de elegir qué alumnos tener en cuenta para un curso concreto, pero esto puede resultar engañoso debido a que en 2º de Bachillerato existe la posibilidad de cursar sólo aquellas materias que se han suspendido y que no han permitido promocionar y obtener el título de bachiller, de modo que se hace necesario tener en cuenta además que el número de asignaturas contempladas a la hora de realizar la media sea mayor o igual que 3, que es el número de asignaturas con el que se debe repetir el curso completo. Con esto sí podemos decir que alumnos que en el último año cursado presentan una calificación media inferior a 2.5 y un número de asignaturas cursadas mayor o igual que 3 son alumnos que han abandonado los estudios en un 95% de los casos.

La cuestión está en si debemos o no eliminar estos alumnos del conjunto de datos a la hora de realizar predicciones. Desde mi punto de vista, creo que todos los cursos anteriores al que abandona son muy válidos a la hora de realizar una predicción, es más, en la práctica hay una considerable parte de alumnado que en ningún momento abandonan los estudios y cuyo expediente es muy similar a el de aquellos que sí abandonan, por lo que a la hora de realizar estimaciones en cursos anteriores, son individuos a tener en cuenta.

Sin embargo, en el curso en el que se produce el abandono podría ser interesante no contemplar tales fuentes de estimación, porque el grado en el que desvirtúan la predicción a simple vista se considera bastante importante. Tendremos que ver qué nos dicen los resultados.

De todas maneras, no se va a tomar ninguna medida referida a eliminar del cálculo de vecinos a este tipo de alumnos. La razón es simple, comprobar el funcionamiento del algoritmo lo que podríamos considerar el peor de los casos, es decir, sin ningún mecanismo capaz de discriminar la información relevante del ruido producido por calificaciones concretas.

3.6 Parametrización del Algoritmo de Filtrado Colaborativo

Puesto que estamos fundamentándonos en algoritmos de CF basados en memoria, debemos establecer las medidas de similitud que vamos a utilizar y el método para la selección de vecinos.

3.6.1 Algoritmos básicos

Las técnicas a utilizar en los que podríamos considerar algoritmos básicos a evaluar serán:

- **PCC y COS como medidas básicas de similitud**: se van a realizar pruebas con el conjunto de datos tanto para el coeficiente de correlación de Pearson (PCC) como para el vector de similitud (COS).
- **KNN para selección de vecinos**: se va a utilizar el método de los *K vecinos más cercanos*, variando *K* hasta encontrar un valor óptimo.
- **WS y WA para predicciones**: como métodos de predicción básicos se van a utilizar tanto la suma ponderada, como la suma media ajustada.

3.6.2 Extensiones y mejoras

En el transcurso de las pruebas se irán aplicando las siguientes extensiones y/o parámetros de mejora que se han estimado convenientes:

- **Frecuencia inversa,** tanto para PCC como para COS.
- **Factor de relevancia,** tanto para PCC como para COS, optimizando el valor de N de forma empírica.
- **Amplificación de casos,** en las predicciones realizadas mediante WA y WS
- **Filtrado basado en ítem**, repitiendo todas las pruebas aplicadas para CF-U

Debemos tener en cuenta que implícitamente estamos aplicando:

- **Representación dimensional reducida,** al agrupar asignaturas tal y como hemos visto en el punto *3.3*.

4. EVALUACIÓN EXPERIMENTAL

4 EVALUACIÓN EXPERIMENTAL

Antes de empezar con la evaluación experimental debe quedar muy claro que los algoritmos y optimizaciones que aquí resulten más exactos y efectivos pueden no funcionar igual en otro ámbito y/o con otro tipo de datos; es algo característico de los algoritmos de filtrado colaborativo y conocido en toda la literatura que este tipo de algoritmos es muy dependiente del dominio concreto en el que actúan [9, 41], y por esto y por la singularidad de los datos con los que pretendemos trabajar es tan interesante estudiar su impacto.

Siguiendo el proceso de experimentación llevado a cabo en [22, 48] se va a dividir esta sección en las siguientes tareas:

- Conjunto de datos a utilizar en la experimentación
- Establecer métricas de evaluación
- Metodología experimental
- Resultados
- Discusión: adecuación de los experimentos y explicación de resultados

4.1 Conjunto de datos

Para la realización de experimentos se han recogido datos académicos reales de un total de 794 alumnos. Estos alumnos pertenecen a dos Institutos de Educación Secundaria de la provincia de Jaén. Los datos de dichos alumnos se han tratado anónimamente, eliminando datos personales y asignando a cada alumno un número entero como identificador.

Los datos del alumnado corresponden a los períodos académicos comprendidos entre el curso escolar 1998-1999 y el curso escolar 2006-2007, para los niveles de 1º y 2º de Bachillerato.

El número total de calificaciones recogidas es de 13421, repartidas entre el número de alumnos mencionados y un total de 74 asignaturas correspondientes a 1º y 2º de Bachillerato. Exactamente 11 de esas asignaturas son comunes, 32 han sido consideradas de modalidad, y 31 son materias optativas. Decir que el único número de materias variable, quiero decir con esto que puede ser distinto de un Instituto de Enseñanza Secundaria a otro, es el de asignaturas optativas.

De entre el total de alumnos, existen 67 individuos para los cuales se puede considerar que en su último año académico cursado abandonan los estudios. El 92,25% de estos alumnos realmente abandonó los estudios en el último curso. El resto continuó para obtener resultados similares al curso siguiente y terminar por dejar el instituto, por lo que hay casos en los que la media obtenida para el curso es inferior a 2.5 para cursos intermedios.

Debemos recordar que aunque en el sistema se realicen predicciones de calificaciones, el objetivo del mismo no es tratar esas predicciones como tales, sino utilizarlas para realizar recomendaciones sobre asignaturas optativas a escoger, y posibles problemas que se puedan plantear en otras materias que forzosamente el alumno deba cursar a lo largo de su devenir académico.

Para terminar con los datos que se van a utilizar, decir que todos los cálculos son *offline*, el usuario no introduce activamente ningún dato, sino que éstos ya han sido almacenados por el personal competente en su momento y al sistema simplemente se le solicita una recomendación para un alumno en concreto. Este factor repercutirá favorablemente en la rapidez del proceso a la hora de dar una recomendación, en lo que al tiempo que se debe emplear desde que el usuario empieza a solicitarla hasta que se obtiene se refiere.

4.2 Métricas de evaluación

En [9] podemos encontrar un amplio estudio de las medidas de evaluación para estos sistemas y que son las más utilizadas en casi la totalidad de la literatura consultada. En nuestro caso vamos a considerar las siguientes:

4.2.1 Precisión

El error absoluto medio (en adelante MAE - *Mean Absolute Error*) está considerado como una medida estadística para la estimación de la exactitud con la que el sistema realizará las predicciones. Este tipo de medidas tratan de verificar con qué grado de exactitud el sistema predice las valoraciones con respecto a las verdaderas valoraciones del usuario.

En nuestro dominio particular, esta medida lo que nos va a decir es cuán lejos están las calificaciones que el sistema predice de las calificaciones que realmente ha obtenido el alumnado. Visto de otro modo, el MAE nos va a decir la habilidad que nuestro algoritmo CF va a presentar a la hora de predecir las calificaciones de un alumno dado su expediente y aquellas recogidas para otros alumnos.

No debemos perder de vista de que esta medida, como su propio nombre indica, es una *media* del error que se produce, lo que quiere decir que aunque el valor de ese error absoluto medio sea bajo, pueden existir alumnos para los que la predicción es errónea completamente.

$$MAE = \left|\overline{E}\right| = \frac{\sum_{i=1}^{n} \left| p_i - r_i \right|}{P}$$

(Ecuación 9, MAE)

Considerando P el número total de predicciones realizadas, p_i el valor de la predicción para una calificación, y r_i el valor real que el alumno obtuvo en esa calificación, podemos ver la forma de calcular el error absoluto medio en la Ecuación 9. Cuanto menor sea el error absoluto medio, mayor será la exactitud con la que el sistema realizará las predicciones.

Es sabido que esta medida no es la más adecuada cuando en el sistema la salida que se ofrece es una lista ordenada con las posibles elecciones a realizar, puesto que el usuario en la mayoría de los casos elige las que se encuentran en la parte más alta de dicha lista. Sin embargo, en nuestro caso es muy importante medir inicialmente el grado de error que vamos a tener en la predicción de las calificaciones por dos razones: primero, estamos tratando con alumnos, por lo que solicitar el máximo de exactitud posible puede considerarse lo mínimo a requerir; segundo, el propósito de este trabajo también es el de evaluar qué otros resultados pueden derivarse de la experimentación que realicemos y para qué otras tareas puede aplicarse este algoritmo.

La granularidad de las valoraciones sabemos que es bastante grande, si la comparamos con la mayoría de los sistemas comerciales en uso que utilizan escalas binarias, de 1-5 o de 1-7, mientras que nuestra escala contiene 11 valores que van de 0 a 10, Por ello resulta muy interesante comprobar qué resultado nos aporta el MAE.

4.2.2 Cobertura

La cobertura es una medida del porcentaje de ítems para los que el sistema puede proporcionar una predicción, en nuestro caso, el porcentaje de asignaturas para las que se solicita predicción y realmente se ha podido estimar la calificación. Esta medida tiene mucho sentido en los algoritmos CF puesto que el propio método de selección de vecinos puede provocar que de entre los K vecinos seleccionados no exista ninguno con valoración para el ítem que deseamos obtener una predicción.

Por otro lado, la cobertura nos proporciona el grado de confianza que nos aporta el MAE, puesto que un MAE de 0, cuando de 20 predicciones sólo se han podido realizar 2 estimaciones (una cobertura del 10%) no ofrece la misma seguridad que un MAE de 1,5 para un 95% de predicciones. Es por esto por lo que la cobertura debe ser medida a la misma vez que la precisión, para que pueda ponerse a punto de forma que no se obtenga beneficio solamente para una de las dos medidas.

Como podemos deducir, un sistema con una baja cobertura será poco valioso de cara a cubrir las necesidades de cualquier usuario medio, puesto se limitaría enormemente el espectro de recomendaciones o alternativas que el sistema propondría a la hora de tomar una decisión.

La forma más fácil de medir la cobertura es elegir un número aleatorio de pares alumno/asignatura, realizar las predicciones del sistema para esos pares en

concreto, y medir el porcentaje de pares para los que se ha podido realizar una predicción. Nosotros calcularemos la cobertura sólo para las asignaturas del curso que se solicite, es decir, si se solicita recomendación para la matrícula de 2º de Bachillerato, se medirá el porcentaje de asignaturas de este curso para el que se puede proporcionar predicción.

4.3 Metodología experimental

Nuestro objetivo es evaluar por separado el funcionamiento del CF-U (filtrado colaborativo basado en usuario) y del CF-I (filtrado colaborativo basado en ítem), tratando de averiguar qué medida de similitud funciona mejor a la hora de elegir usuarios, y cuál de predicción a la hora de estimar las calificaciones para cada uno de ellos.

Vamos a dividir las pruebas y los resultados experimentales en 3 subconjuntos:

a) resultados obtenidos para algoritmos colaborativos basados en usuario

b) resultados obtenidos para algoritmos colaborativos basados en ítem

c) comparación entre resultados basados en usuario y basados en ítem

Debido a la gran cantidad de experimentos realizados (tanto en número como en variantes de los algoritmos), cada subconjunto de los anteriores se va a subdividir en dos tipos de experimentos:

a) experimentos cuya finalidad es apuntar de forma general cuál o cuáles de entre las distintas variantes del algoritmo colaborativo empleado son las que mejor se comportan a priori

b) experimentos para decidir el mejor algoritmo y la optimización de sus posibles parámetros, como por ejemplo, el número de vecinos que se utilizará para las predicciones

4.3.1 Algoritmos y variantes a contemplar contemplar

Aunque ya se mencionó anteriormente, con respecto a los alumnos que supuestamente abandonan los cursos, inicialmente se van a realizar pruebas teniéndolos en cuenta, de modo que los resultados que se obtendrán serán para el peor de los casos, es decir, cuando este tipo de alumnos afectará negativamente al proceso de selección de vecinos y de predicción de calificaciones. En la parte correspondiente a decidir las variantes que en un principio mejor se comportan se van a realizar pruebas para distintas medidas de similitud y consiguientes extensiones y mejoras.

Con respecto al cálculo de vecinos, se va a usar tanto el Coeficiente de correlación de Pearson como el Vector de Similitud (Distancia del Coseno), con las siguientes mejoras y/o extensiones:

- Aplicación de la frecuencia inversa a PCC y COS.
- Aplicación de factor de relevancia para PCC y COS.
- Aplicación mixta de los dos parámetros anteriores a PCC y COS.
- Variante de PCC usando como medias de los alumnos un valor fijado a 5.

Inicialmente, a la hora de aplicar el factor de relevancia, no se va a contemplar un umbral de asignaturas coincidentes, es decir, se hará de forma que quien más asignaturas tenga en común tendrá más relevancia frente a otros que compartan menos. Esto se conseguirá multiplicando directamente la similitud obtenida por el número de asignaturas comunes. Debemos destacar que de esta forma el PCC deja de tomar valores entre -1 y 1, aunque este hecho no afecta al cálculo general. Posteriormente se optimizará este parámetro de forma empírica para calcular de qué manera se obtienen los mejores resultados posibles.

Otra de las modificaciones que se va a realizar consiste en fijar la media a la hora del cálculo de similitudes en el coeficiente de correlación de Pearson. Si observamos este coeficiente, las medias que utiliza son relativas a los usuarios concretos, con lo cual la estimación se produce de una forma relativa también a esas medias; podemos intentar hacer que el cálculo de la similitud sea absoluto, si en vez de utilizar como media la del alumno, utilizamos una media común, la media de valoración por asignatura, es decir, 5; de esta manera no se considerará similar un alumno que presenta por ejemplo calificaciones de (5,4,6,5) frente a otro que tenga las calificaciones (9,8,10,9), hecho que sí ocurría usando la media de cada alumno.

Por otro lado, para el cálculo de predicciones se van a realizar pruebas para la suma ponderada y suma media ajustada, y posteriormente se aplicará la variante de la amplificación de casos tanto para WS como para WA.

4.3.2 Procedimiento de cada iteración

Cada iteración conlleva una serie de pasos que vamos a enumerar a continuación:

1. Inicialmente se establece el porcentaje para el conjunto de prueba al 10%.
2. Mientras el porcentaje del conjunto de prueba sea <= 30%
    a. se establecen el conjunto de prueba y el conjunto de entrenamiento
    b. Mientras queden medidas de similitud por probar para el par prueba-entrenamiento
        i. Para la medida elegida se calcula la similitud entre vecinos para el conjunto de prueba

ii. Se realizan predicciones con WA y WS tomando *K* (número de vecinos a contemplar) distintos valores: 10, 20 y 30

iii. Se calculan los valores para el MAE y la cobertura y se almacenan los resultados

c. Se aumenta en 10 el porcentaje del conjunto de prueba

La formación de los conjuntos de prueba y entrenamiento en función del porcentaje para el conjunto de prueba establecido, que tomará los valores 10, 20 y 30, se realiza de la siguiente manera:

- Conjunto de prueba y conjunto de entrenamiento para CF-U: se extraen del conjunto de datos un porcentaje de asignaturas de forma aleatoria, de manera que no se tengan en cuenta para el cálculo de vecinos entre alumnos y se realizan las predicciones para todos los alumnos que tengan calificaciones en esas asignaturas, con lo que podremos aplicar las métricas de evaluación. Por ejemplo, si el porcentaje establecido es del 10%, un 90% aleatorio de las asignaturas formarán parte del conjunto de entrenamiento, y el 10% restante pasarán al conjunto de prueba.
- Conjunto de prueba y conjunto de entrenamiento para CF-I: se extraen del conjunto de datos un porcentaje de alumnos de forma aleatoria, de manera que no se tengan en cuenta para el cálculo de vecinos entre asignaturas y se realizan las predicciones para todas las materias que tengan esos alumnos, para poder aplicar posteriormente las métricas de evaluación. Por ejemplo, si el porcentaje establecido es del 30%, un 70% aleatorio de los alumnos formarán parte del conjunto de entrenamiento, y el 30% restante pasarán al conjunto de prueba.

4.3.3 Optimización de parámetros

Terminado el proceso anterior, se realizarán nuevas pruebas para ajustar aquellos parámetros y mejoras que hayan dado mejores resultados, para afinar aún más a la hora de escoger qué parámetros concretos utilizar. Decididas la o las medidas de similitud que mejor funcionan, se tratarán de optimizar variando el valor de *K* y también el de *N* en el caso de que se utilice el factor de relevancia.

En el caso de *K*, y siguiendo el procedimiento explicado para cara iteración, se tomarán valores de *K* entre 5 y 50, utilizando todos los múltiplos de 5 del intervalo, y usando sólo las medidas candidatas escogidas en función de los resultados.

Optimizados los parámetros se realizará un último paso consistente en alterar la forma en la que se escogen las materias para el conjunto de prueba, de

modo que en la medida de lo posible (depende del porcentaje de prueba) sean todas de una única tipología, aunque dentro de dicha tipología se seguirían escogiendo de forma aleatoria: o bien optativas, o bien de modalidad, o bien comunes. Debido a su escaso número, en el caso de las asignaturas comunes es necesario rellenar el conjunto de prueba con materias de los otros dos tipos, proceso que se seguirá haciendo de forma aleatoria como hasta el momento.

El número mínimo de iteraciones realizadas para las medidas de similitud será de 50, y si no se expresa lo contrario se mostrarán los resultados agrupados puesto que hasta ese momento lo único que nos interesará será comprobar el funcionamiento general de las distintas alternativas.

Un último apunte a realizar tiene que ver con los valores que tomarán el MAE y la cobertura. El error absoluto medio, en nuestro caso, puede tomar valores que van desde 0 hasta 10, Un MAE de 0 indicaría que las predicciones coincidieron perfectamente con los resultados obtenidos. Por ejemplo, podríamos obtener un error medio de 3 para un alumno con las siguientes predicciones y calificaciones reales (Tabla 4):

		Predicción	Real
Materia	M1	4	7
	M2	2	5
	M3	8	5
	M4	10	7
	M5	9	9
	M6	0	6

TABLA 4. Comparación entre predicción y calificación real

Como podemos suponer, un error absoluto medio de 3 no es nada deseable. En la literatura relativa a sistemas de filtrado colaborativo se consideran valores adecuados de MAE aquellos que rondan entre el 0,6 (0,15 arriba o abajo) para valoraciones entre 1 y 5, y los que se mueven en torno al 0,7 (0,15 arriba o abajo) con escalas de valoraciones entre 1 y 7, si lo expresamos en porcentaje, se predice la valoración con entre un 10 y un 14% de error. En nuestro dominio un margen de error como este sería relativamente aceptable, y si los resultados estuvieran en ese intervalo o inferior sería posible plantearse la construcción de un sistema como el que hemos ido dibujando hasta ahora.

4.4 Resultado experimental

Es momento de presentar los resultados experimentales obtenidos aplicando el filtrado colaborativo al problema que hemos planteado, y con las

decisiones tomadas explicadas a lo largo de este documento. La Tabla 5 define las abreviaturas que se van a utilizar para las etiquetas de los resultados:

MEDIDAS DE SIMILITUD	COEFICIENTE DE CORRELACIÓN DE PEARSON			
	pcc	Coeficiente de correlación de Pearson utilizando 5 como media prefijada.	**pcr**	Coeficiente de correlación de Pearson utilizando como media la de cada alumno concreto.
	pcn	Variante de pcc a la que se multiplica el número de asignaturas comunes	**prn**	Variante de prn a la que se multiplica el número de asignaturas comunes
	pci	Variante de pcc a la que se aplica la frecuencia inversa	**pni**	Variante de pcc a la que se aplica tanto la frecuencia inversa como el multiplicar por el número de asignaturas comunes
	VECTOR DE SIMILITUD (COSENO)			
	cos	Medida del Coseno basada en Vector de Similitud	**coi**	Variante de cos aplicando la frecuencia inversa
	con	Variante de cos multiplicando por el número de asignaturas comunes	**cni**	Variante de cos a la que se aplica tanto la frecuencia inversa como el multiplicar por el número de asignaturas comunes

MÉTRICAS	MÉTRICAS DE EVALUACIÓN	
	wa	Predicciones realizadas mediante *suma media ajustada*
	ws	Predicciones realizadas mediante *suma ponderada*
	waca	Predicciones realizadas aplicando a la suma media ajustada la amplificación de casos
	wsca	Predicciones realizadas aplicando a la suma ponderada la amplificación de casos
	Cob	Cobertura: porcentaje de materias para las que se debía realizar una predicción y efectivamente se ha realizado esa predicción

TABLA 5. Abreviaturas y terminología usada en las pruebas experimentales

4.4.1 Resultados obtenidos para CF-U

Empezaremos con los datos obtenidos para CF-U inicialmente usando sólo como medidas de predicción WA y WS (sin amplificación de casos).

Resultados generales

En la Tabla 6 y la Figuras 13 y 14 podemos ver el error absoluto medio y la cobertura para resultados generales agrupados obtenidos para las distintas medidas de similitud y las dos técnicas de predicción básicas:

RESULTADOS GENERALES PARA CF-U			
Medida de Similitud	MAE		Cobertura (en %)
	WA	WS	
cni	1,3023	1,7383	79,9063
coi	1,2054	1,5703	88,3686
con	1,1913	1,8717	97,8888
cos	1,3084	1,808	69,5259
pcc	1,1266	1,1322	82,9623
pcr	1,1792	1,6138	78,3836
pci	1,5398	2,7062	68,8913
pcn	**0,9523**	1,0327	**98,6686**
prn	1,0286	1,4437	97,0385
pni	1,4839	2,8807	71,2228

Tabla 6. Resultados generales para CF-U

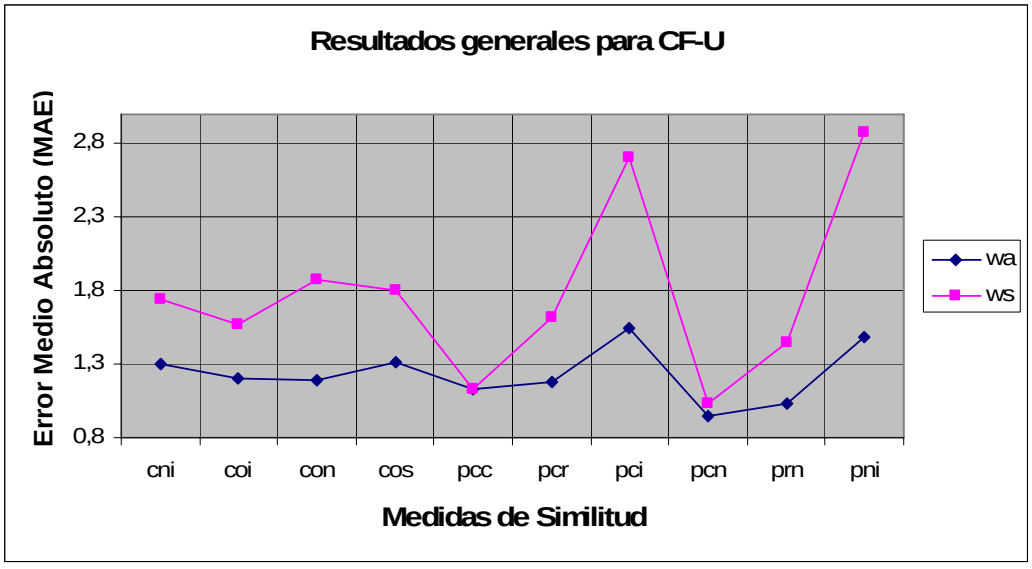


Figura 13. Resultados generales para CF-U

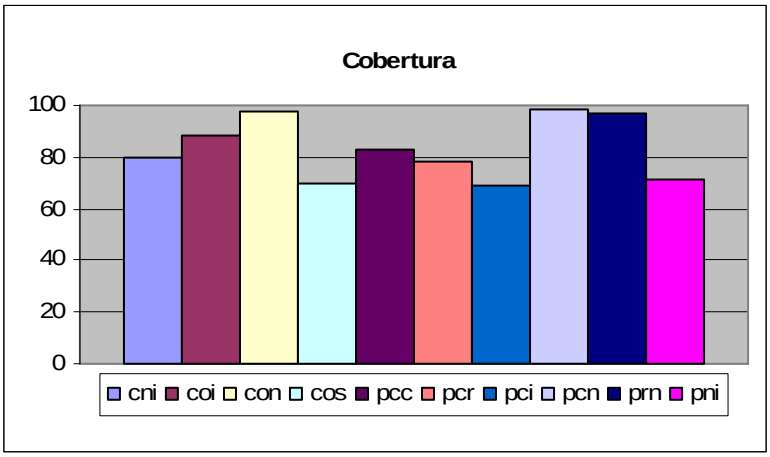


Figura 14. Coberturas para distintas aproximaciones de CF-U

Como hemos indicado anteriormente, estos resultados han sido agrupados para las distintas iteraciones realizadas, valores de *K*, tamaños para el par de conjuntos entrenamiento-test, etc.. Más adelante se mostrarán resultados más desglosados; por ahora, para un propósito más general estos datos nos sirven.

De lo que hemos observado en los gráficos anteriores y en la tabla de resultados podemos derivar varias conclusiones:

- El método de predicción que a priori mejor funciona es el de la suma media ajustada. Para todas y cada una de las medidas de similitud utilizadas, la suma media ajustada funciona considerablemente mejor que la suma ponderada.
- Como medida de similitud general, el coeficiente de correlación de Pearson obtiene un MAE mucho menor que el vector de similitud, excepto cuando tratamos de utilizar la frecuencia inversa. Esto es debido a que dicha frecuencia inversa es una mejora pensada especialmente para el vector de similitud, y utilizada sobre todo en recuperación de información, por lo que la adaptación a su uso con PCC no tiene por qué ofrecer mejores resultados.
- Con respecto al coseno, la aplicación de los parámetros de mejora de forma individual (frecuencia inversa y factor de relevancia) consigue un beneficio en el MAE significativo, aunque no hasta el punto de superar los resultados del PCC.
- Para este PCC, el único parámetro de mejora que no repercute favorablemente en los resultados es, como hemos mencionado anteriormente, el de la frecuencia inversa. El del factor de relevancia aporta un beneficio en torno al 0,1, suceso bastante deseable.

Un suceso a destacar es el hecho de que el uso de una media fija en el cálculo de similitudes para PCC funcione considerablemente mejor que el usar la propia media de los alumnos. Sin embargo, si analizamos el dominio en el que nos movemos y por qué en otros dominios resulta beneficioso el utilizar la media relativa, encontraremos la explicación.

En los sistemas comerciales, el utilizar la media de cada usuario viene motivado porque se asume que existirán usuarios que aun usando criterios parecidos de valoración, unos pueden tener la costumbre de votar siempre con valores altos, otros con valores medios y otros con bajos. El uso de esta media disminuye el efecto producido por estos comportamientos y se beneficia de ellos. Sin embargo, en nuestro dominio los alumnos no califican las asignaturas en base a su gusto, sino que obtienen las calificaciones debido a que presentan unas buenas aptitudes/actitudes hacia una materia determinada, se han esforzado, y en cierta medida, puede que porque les ha gustado. Esto puede provocar calificaciones muy dispares. Si bien es verdad que existen alumnos que siempre tienden a sacar buenas notas, y otros que tienden a suspender, parece ser que su existencia no repercute en el hecho de que medir la similitud de una forma absoluta sea mejor que hacerlo de forma relativa.

En definitiva, y como hemos visto en los resultados, el mejor error absoluto obtenido, coincidiendo con la mejor cobertura (hecho que le aporta mayor relevancia a este dato), corresponde a la medida de similitud que calcula el coeficiente de correlación de Pearson pero usando una media fija de 5, y calculando las predicciones en base a la suma media ajustada, obteniéndose un error medio de 0,9523 y una cobertura del 98,6686%.

¿Podemos decir que un error medio de 0,9523 es aceptable? Pues bien, teniendo en cuenta que estamos manejando una escala de de 0 a 10, y no perdiendo de vista el objetivo final que se persigue (recomendar asignaturas optativas, e incluso avisar en qué asignaturas puede necesitarse refuerzo), tal error es, desde mi punto de vista, aceptable. Estamos tratando con un **error en torno al 10%**, y antes mencionamos un error de entre el 10 y el 14% como el que se suele dar por aceptable en las aplicaciones comerciales.

Con un sistema de recomendación basado en estas técnicas, un error de 1 muy difícilmente recomendaría una asignatura comprometida (ya nos encargaríamos de establecer el rango de valores adecuado para la recomendación); por eso volvemos a decir que un error de 1 es aceptable.

Se han realizado otra serie de pruebas para esclarecer la siguiente duda: ¿debemos realizar las predicciones teniendo en cuenta decimales, cuando en las calificaciones reales no se aportan? Si es así, ¿cuántos decimales serían adecuados?

Para ello, y para tener una idea más clara de la aproximación a la hora de medir la similitud que mejor resultado nos aporta, se han realizado una serie de

pruebas usando *pcn* y *prn* (las dos mejores medidas de similitud según los resultados anteriores), utilizando distintos números de decimales en las predicciones: desde 0 hasta 4 decimales.

Igual que en el caso anterior, los resultados que se van a mostrar están agrupados y comprenden los valores de *K* y los distintos conjuntos utilizados de entrenamiento-test.

Predicciones con distinto número de decimales				
Medida de Similitud	Nº de decimales	MAE		Cobertura
		wa	ws	
pcn	**0**	**0,942**	1,0161	98,7995
pcn	1	0,9724	1,0428	
pcn	2	0,9727	1,0429	
pcn	3	0,9602	1,0366	
pcn	4	0,9602	1,0366	
prn	0	1,0342	1,4401	97,3939
prn	1	1,0614	1,4575	
prn	2	1,0617	1,4577	
prn	3	1,0421	1,4589	
prn	4	1,0421	1,4589	

Tabla 7. Uso de distinto número de decimales

Tanto en el caso de *pcn* como en el de *prn* podemos comprobar que aunque aumentemos el número de decimales, los mejores resultados se producen sin usar decimales, usando números enteros y redondeando las predicciones que obtenemos.

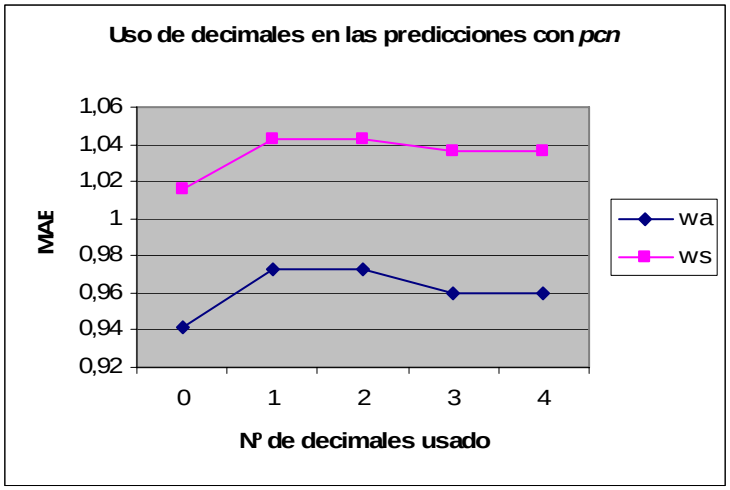


Figura 15. Gráfica sobre el uso de decimales en CF-U

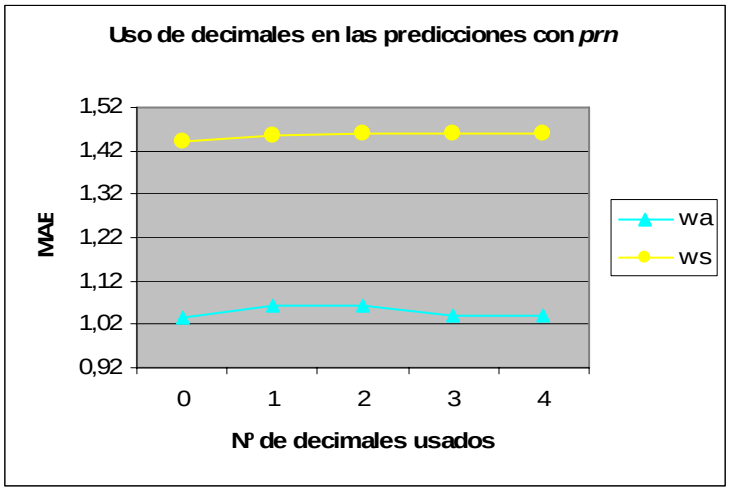


Figura 16. Gráfica sobre el uso de decimales en CF-U

Optimización de parámetros

Ahora vamos a intentar poner a punto los parámetros para conseguir la mejor precisión posible. Empezaremos por estimar el mejor valor para el número de vecinos con el que vamos realizar las predicciones, *K*. Para ello vamos a realizar pruebas con valores de *K* comprendidos entre 15 y 35, incrementándolos de 5 en 5. No se tomarán inicialmente valores más bajos ni más altos, a no ser que los datos demuestren que debe hacerse.

Las pruebas se realizarán para las medidas de similitud que mejores resultados nos han dado, es decir, *pcn* y *prn*, centrándonos en mejorar el funcionamiento de ambos.

Predicciones para distintos valores de K				
Similitud	Valor de K	MAE		Cobertura
		wa	waca	
pcn	15	0,9346	0,9331	98,7347
pcn	20	0,9289	0,9271	99,0915
pcn	25	0,9271	0,9242	99,2857
pcn	30	0,9272	0,9237	99,4072
pcn	35	0,9262	0,9224	99,4509
prn	15	0,9926	0,999	97,0256
prn	20	0,9823	0,9864	97,5053
prn	25	0,9763	0,9795	97,7733
prn	30	0,9726	0,9762	97,9142
prn	35	0,9993	1,0047	97,8824

Tabla 8. Resultados para distintos valores de K

Cabe notar que hemos eliminado de los resultados el MAE obtenido por WS, debido a que como hemos ido viendo hasta ahora WA aporta mejores predicciones, y por otro lado hemos incorporado un nuevo parámetro de mejora, esta vez en las predicciones: se trata de la amplificación de casos, utilizada tal cual vimos en el punto 2 para las estimaciones realizadas mediante WA.

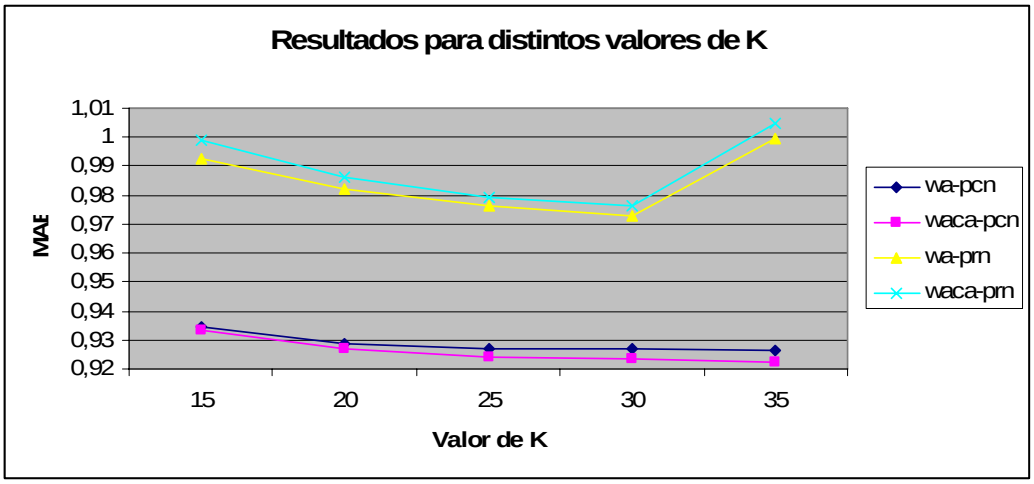


Figura 17. Gráfica sobre el uso de distintos valores de K en CF-U

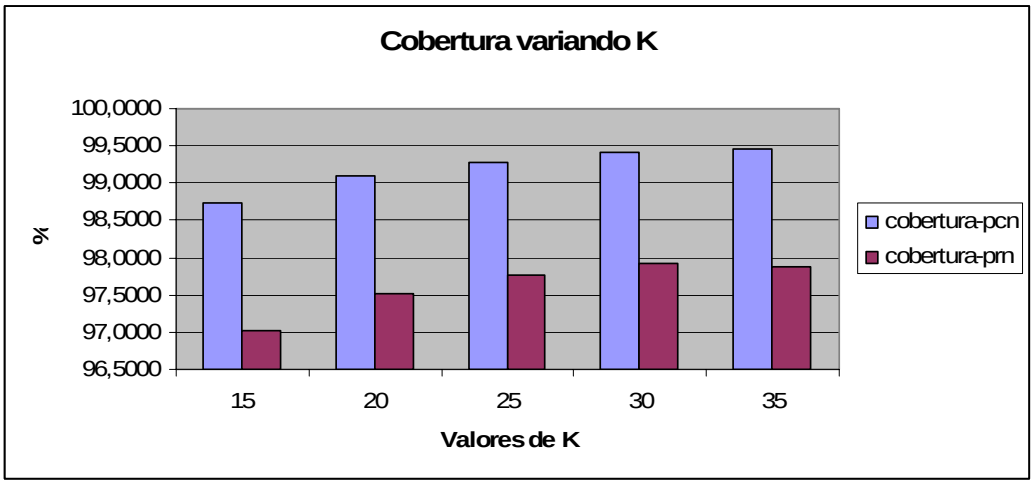


Figura 18. Gráfica sobre el uso de distintos valores de K en CF-U

Definitivamente, y para todos los valores de *K*, la medida de similitud que mejor funciona tanto en exactitud como en cobertura es *pcn*, es decir, el uso del coeficiente de correlación de Pearson con la utilización de media fija de valor 5. Esto nos hace rechazar el resto. Por otro lado, vemos que aunque sea ligeramente, para el caso de *pcn* la amplificación de casos mejora los resultados en todos los valores de *K*.

Con respecto a éste parámetro *K*, vemos que conforme va aumentando, va mejorando la precisión, por lo que vamos a hacer un estudio más concienzudo, realizando nuevas pruebas y tomando un rango más amplio para *K*.

Precisión para distintos valores de K			
Similitud	Valor de K	MAE waca	Cobertura
pcn	5	0,9764	95,5096
pcn	10	0,949	97,8447
pcn	15	0,9331	98,7347
pcn	20	0,9272	99,1158
pcn	25	0,9242	99,3119
pcn	30	0,9235	99,4323
pcn	35	0,923	99,4955
pcn	40	0,9262	99,5892
pcn	45	0,9292	99,595
pcn	50	0,9311	99,6275

Tabla 9. Precisión para distintos valores de K

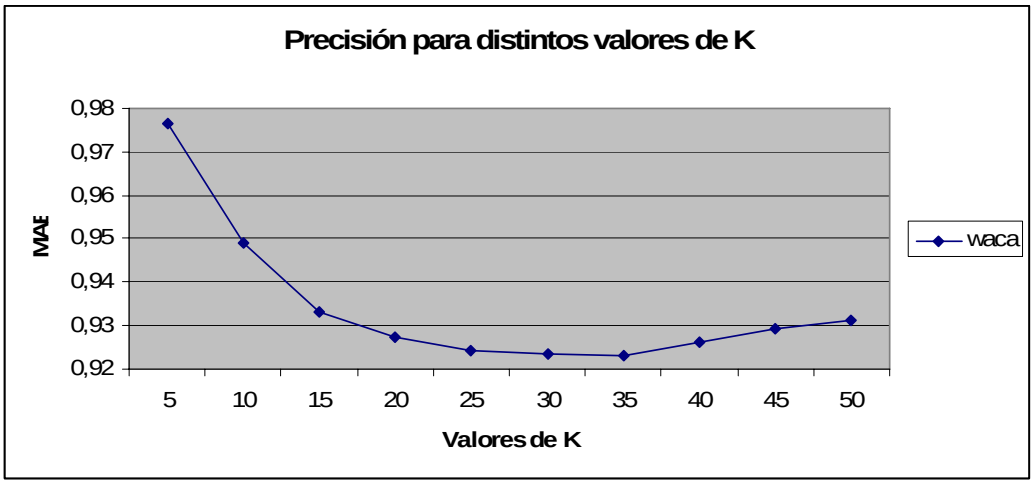


Figura 19. Gráfica sobre el uso de distintos valores de K en CF-U

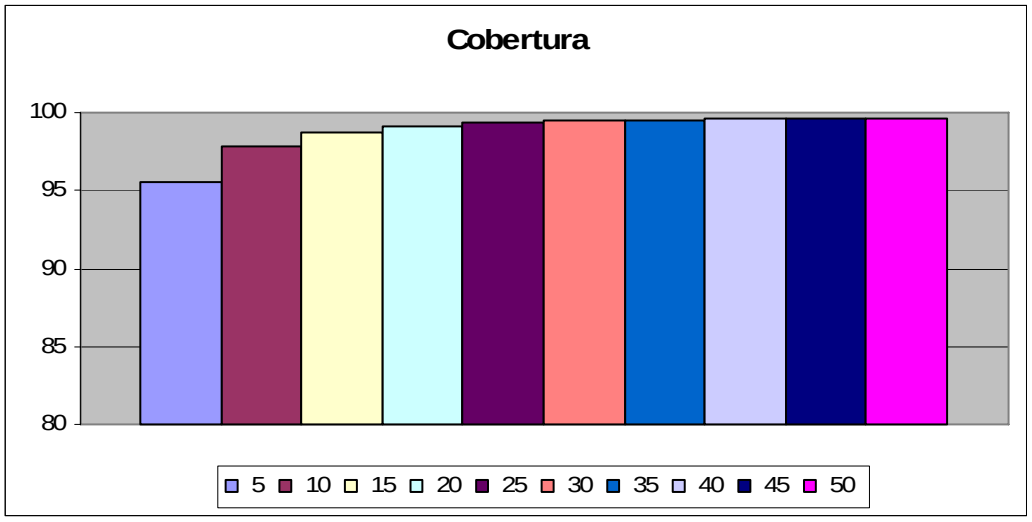


Figura 20. Cobertura con distintos valores de K en CF-U

Como podemos comprobar, los cambios en la cobertura empiezan a no ser significativos a partir de un valor para *K* de 20 o 25, aunque se sigue mejorando. Sin embargo, para el caso de la precisión podemos observar que el MAE mejora hasta que *K* toma un valor de 35, y a partir de ahí empeora ligeramente. En el rango de 20 a 40 el error medio absoluto ronda en torno a un 0,92, y la diferencia es menos apreciable si nos ajustamos a los valores 25, 30 y 35. Cualquiera de estos valores puede ser una decisión acertada a la hora de asignar el parámetro *K*.

El último proceso de optimización que vamos a tener en cuenta para el algoritmo CF-U es el de encontrar el mejor valor para N, refiriéndonos al parámetro utilizado en el factor de relevancia. Probaremos distintos valores, hasta comprobar qué número de asignaturas compartidas empieza a dejar de ser significativo. Las pruebas se realizarán para valores de *K* entre 25 y 35, tras ver los resultados anteriores.

Como podemos comprobar en la tabla y en el gráfico posterior, el mejor resultado se obtiene dividiendo el número de asignaturas compartidas entre los alumnos por 50, como ya se proponía en [8], mejorando los resultados que obteníamos hasta el momento al tener en cuenta todas las asignaturas compartidas.

En la mayoría de las pruebas observadas con los distintos valores de N, el valor de *K* que mejor ha funcionado ha sido 30, excepto para N=10, N=15 y para N no definido; para el resto de pruebas un *K* con valor de 25 o 35 obtiene peores resultados.

Precisión para distintos valores de N			
Valor de N	MAE para waca		
	K=25	K=30	K=35
10	0,9408	0,937	0,935
15	0,9279	0,9283	0,9293
20	0,9216	0,9204	0,921
25	0,9298	0,9271	0,9287
30	0,9279	0,9252	0,9264
35	0,9257	0,9238	0,9246
40	0,9216	0,9199	0,9209
45	0,9216	0,9199	0,921
50	0,9187	**0,9172**	0,9181
N	0,9242	0,9235	0,923

Tabla 10. Uso de distintos valores de N en el factor de relevancia

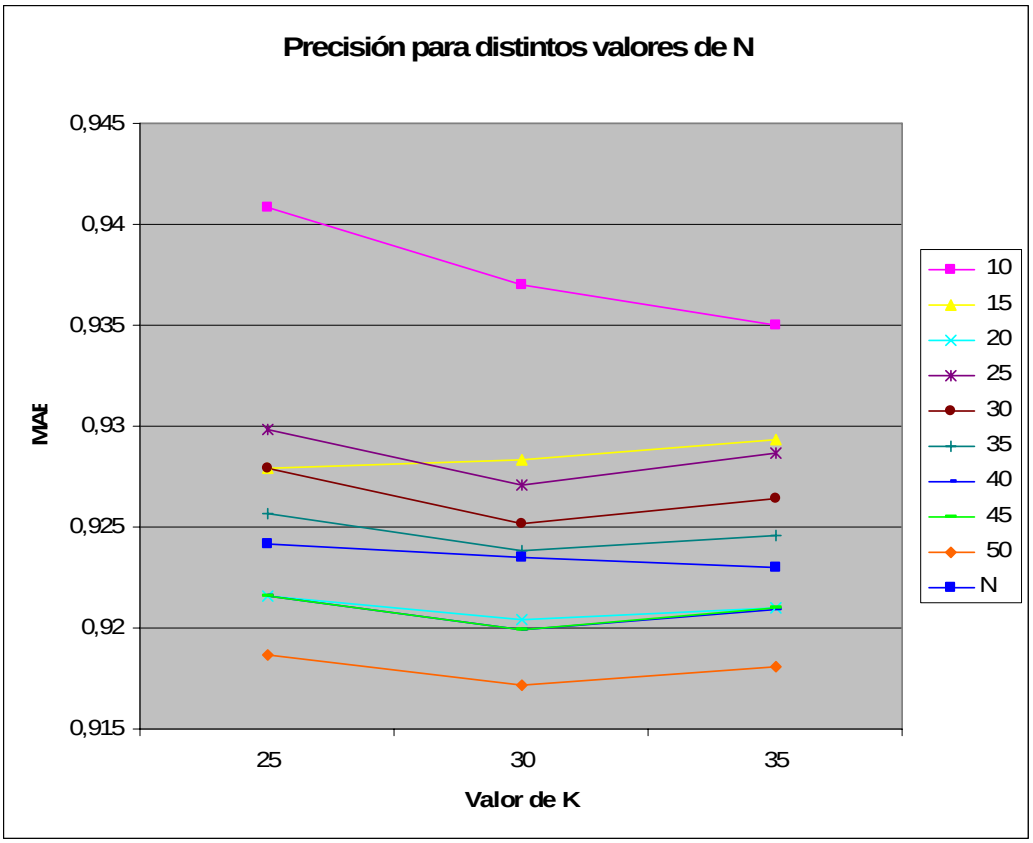


Figura 21. Gráfica sobre el uso de distintos valores de N en el factor de relevancia

En definitiva, podríamos decir que el algoritmo que mejor se ha comportado en nuestro conjunto de datos sería un *K-NN* con los siguientes parámetros:

- Uso para cálculo de similitud del coeficiente de correlación de Pearson con media fijada a 5
- Aplicación del parámetro de mejora para Pearson correspondiente al factor de relevancia, multiplicando la similitud obtenida por *n/50* si *n<50*, donde *n* es el número de asignaturas compartidas por los alumnos en cuestión.
- Elección de los 30 mejores vecinos para las predicciones.
- Predicciones realizadas redondeando al entero utilizando la suma media ajustada de las predicciones de los vecinos seleccionados.
- Aplicación del parámetro de mejora para la suma media ajustada correspondiente a la amplificación de casos.

Con estos parámetros, y para el total de pruebas realizadas, se ha obtenido un error absoluto medio de 0,9172 (alrededor del 8%) y una cobertura del 99,4705% de las predicciones.

4.4.2 Resultados obtenidos para CF-I

Vamos a ver ahora qué tal se comporta el filtrado colaborativo basado en ítem a la hora de realizar predicciones para nuestro problema concreto.

Resultados generales

Como antes, empezaremos con los datos obtenidos inicialmente usando sólo como medidas de predicción WA y WS (sin amplificación de casos) y agrupados.

RESULTADOS GENERALES PARA CF-I			
Medida de Similitud	MAE		Cobertura
	wa	ws	
cni	1,4558	2,0404	45,3995
coi	1,0603	1,4096	98,4027
con	1,0993	1,4207	**99,9568**
cos	1,1999	1,4895	73,0511
pcc	1,7377	2,3557	82,9186
pci	1,4684	1,9086	69,5996
pcn	1,0817	1,2545	99,7117
pcr	**1,0385**	1,3339	98,3152
pni	1,136	1,4682	98,6596
prn	1,0493	1,3585	99,8764

Tabla 11. Resultados generales para CF-I

En el caso de CF-I también obtiene mejores resultados el uso de la medida para similitudes basada en el coeficiente de correlación de Pearson, aunque el error más bajo se consigue para el uso de Pearson con la media relativa de las asignaturas. De todas maneras, en este caso las diferencias no son nada esclarecedoras, pues el orden en el que andan es para 5 de las medidas de similitud menor a 0,6, incluyendo en este rango los resultados obtenidos por la similitud basada en el coseno normal, que mejora en 0,03 aplicándole la frecuencia inversa.

A la hora de la predicción sigue funcionando sustancialmente mejor el uso de la suma media ajustada frente a la suma ponderada. El mejor error absoluto obtenido es de 1,0385 (error del 9,44%), correspondiente a Pearson con media relativa de las asignaturas, con una cobertura del 98,3152%, y la mejor cobertura es de 99,9568% obtenida por el uso del coseno y el factor de relevancia, aunque su MAE es de 1,0993. Vemos que el error producido también es aceptable, aunque mayor que en CF-U.

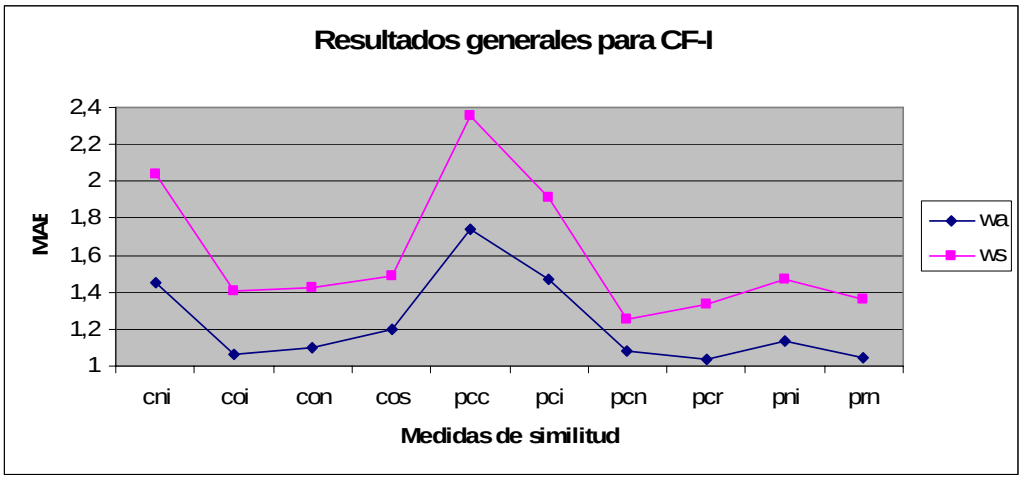


Figura 22. Gráfica sobre resultados generales para CF-I

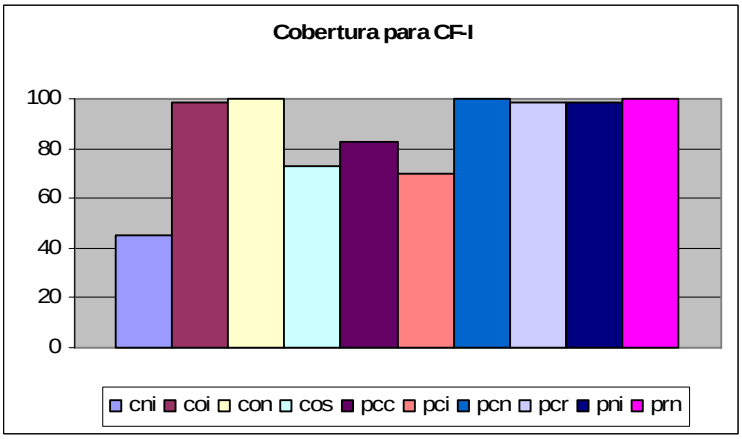


Figura 23. Coberturas para CF-I

Con respecto al uso de decimales en las predicciones se han realizado pruebas similares a CF-U, pero en este caso se ha tratado de clarificar un poco más qué medida obtiene mejores resultados, para lo que se han repetido las pruebas en el caso de las siguientes medidas de similitud:

- Coseno con frecuencia inversa
- Pearson con media relativa a la asignatura
- Pearson con media relativa a la asignatura y factor de relevancia
- Pearson con media fija de 5 y factor de relevancia

Como en el caso de CF-U, hemos usado distintos números de decimales en las predicciones (desde 0 hasta 2 decimales en este caso) obteniendo los resultados de la tabla, en la que sólo aparecen los resultados para la suma media ajustada (WA), puesto que como hemos ido viendo a lo largo de todas las pruebas, presenta siempre unos valores bastante más acertados que la suma ponderada.

Predicciones con distinto número de decimales				
Decimales	MAE			
	coi	pcn	pcr	prn
0	1,054	1,0717	1,0304	1,0432
1	1,1056	1,1058	1,0651	1,072
2	1,1058	1,106	1,0654	1,0723

Tabla 12. Uso de decimales en CF-I

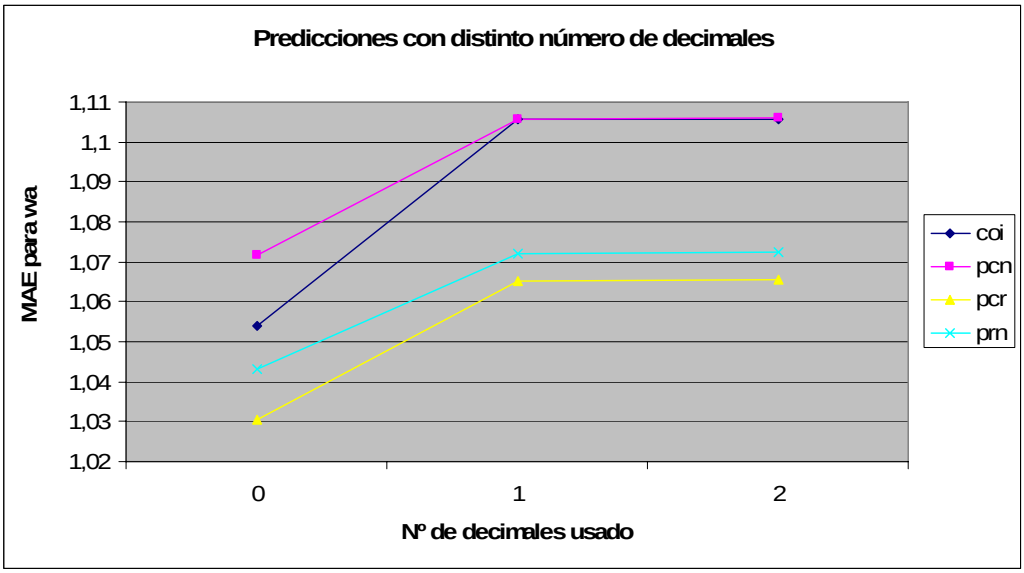


Figura 24. Uso de decimales en CF-I

Podemos comprobar que sigue siendo *pcr* quien muestra los mejores resultados, aunque por escaso margen, y que como en CF-U, el uso de decimales empeora los resultados.

Optimización de parámetros

Ahora vamos a intentar optimizar los parámetros como ya hemos hecho para CF-U de forma que obtengamos la mejor precisión posible. Vamos a empezar por el valor de *K*, y lo vamos a estudiar sobre *pcr* y *prn*.

Predicciones con distintos valores de K				
Similitud	Valor de K	MAE		Cobertura
		wa	waca	
pcr	5	1,1849	1,1871	32,7428
pcr	10	1,0798	1,0581	82,3056
pcr	15	1,0236	1,0188	97,1359
pcr	20	1,0236	**1,0059**	99,602
pcr	25	1,0259	1,0129	99,8382
pcr	30	1,0426	1,0191	99,8721
pcr	35	1,0766	1,0392	99,8996
pcr	40	1,058	1,0955	99,869
pcr	45	1,2174	1,0915	99,9048
pcr	50	1,1728	1,146	99,9048
prn	5	1,0496	1,0404	99,4638
prn	10	1,0603	1,0327	99,8501
prn	15	1,0499	1,0336	99,8731
prn	20	1,0527	1,0327	99,8866
prn	25	1,0505	1,0337	99,908
prn	30	1,0543	1,0336	99,8917
prn	35	1,0527	1,034	99,908
prn	40	1,0431	1,0342	99,8899
prn	45	1,058	1,0354	99,908
prn	50	1,145	1,075	99,908

Tabla 13. Predicciones con distintos valores de K

Análogamente a CF-U, hemos utilizado también el parámetro de amplificación de casos para la suma media ajustada, el cual podemos comprobar que mejora la gran mayoría de los resultados obtenidos.

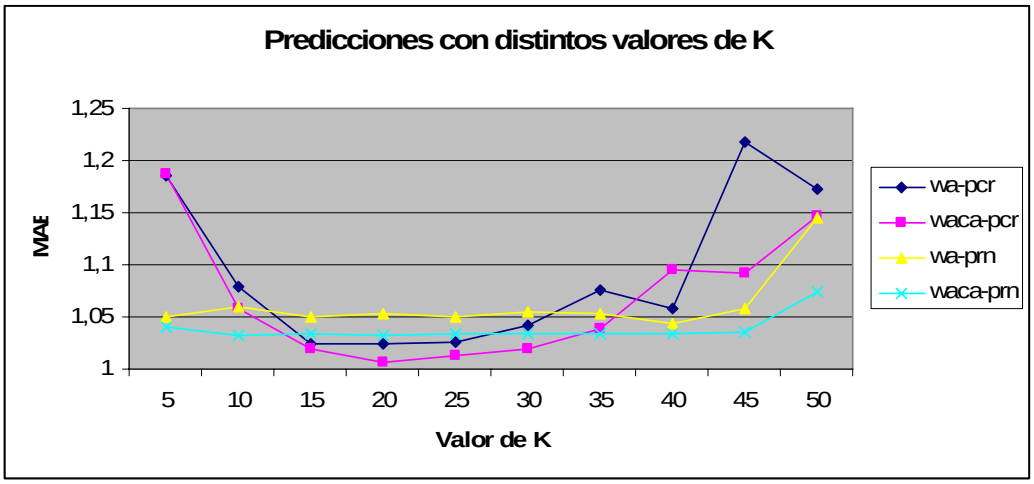


Figura 24. Gráfica de predicciones con distintos valores de K

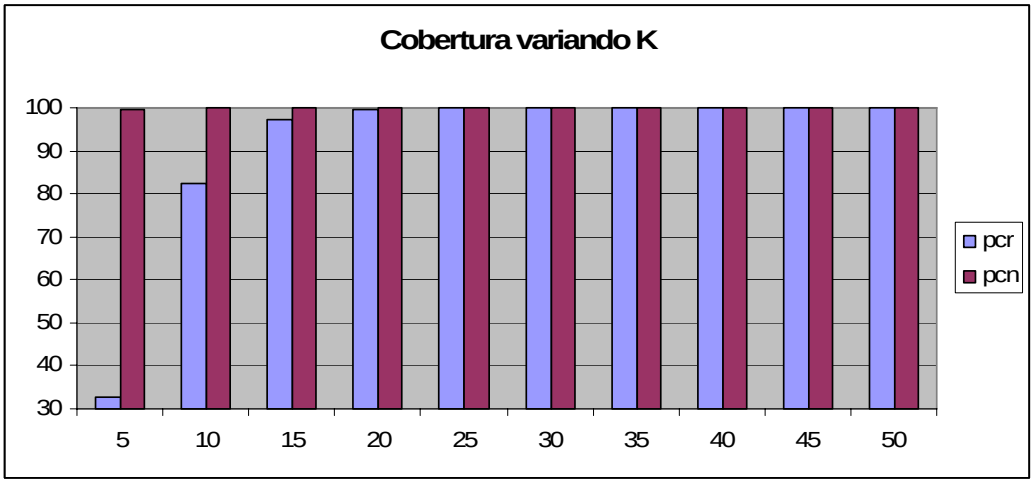


Figura 25. Gráfica de coberturas con distintos valores de K

Con respecto a éste parámetro *K*, vemos que para valores de entre 15 y 25 se obtienen los mejores resultados tanto en *pcr* como en *prn*, encontrándose muy cercanos los resultados obtenidos para esos rangos. Según muestran los resultados en éste caso, el tener el mayor número de alumnos coincidentes posible en cuenta a la hora de calcular la similitud resulta no mejorar el error absoluto medio, por lo que vamos a comprobar si con valores bajos de N se mejoran los resultados.

Como en CF-U, realizaremos pruebas con distintos valores de *K* (15, 20 y 25) y para distintos valores de N aplicando el factor de relevancia a *pcr*.

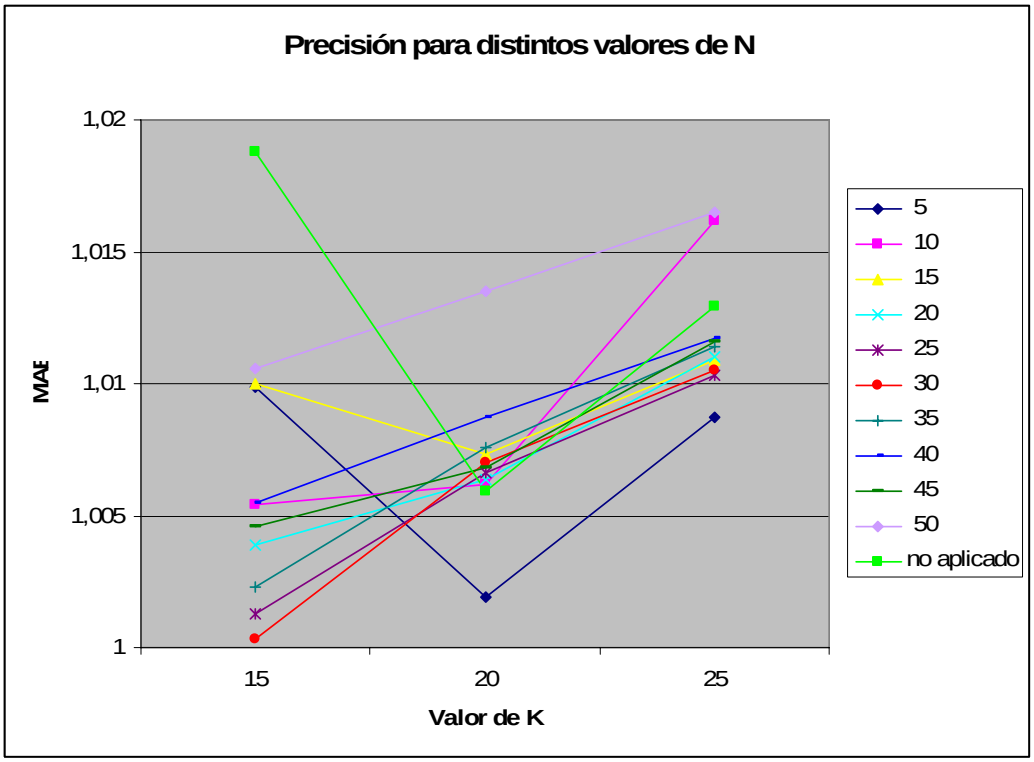


Figura 26. Gráfica de resultados con distintos valores de N

Vemos que el mejor resultado se obtiene dividiendo el número de alumnos coincidentes por 30 para un valor de *K* de 15, mejorando los resultados que se obtenían mediante *pcr* simple.

Precisión para distintos valores de N			
Valor de N	MAE para waca - Valor de K		
	15	20	25
5	1,0099	1,0019	1,0087
10	1,0054	1,0062	1,0162
15	1,01	1,0073	1,0109
20	1,0039	1,0064	1,011
25	1,0013	1,0066	1,0103
30	**1,0003**	1,007	1,0105
35	1,0023	1,0076	1,0114
40	1,0055	1,0087	1,0117
45	1,0046	1,0068	1,0116
50	1,0106	1,0135	1,0165
no aplicado	1,0188	1,0059	1,0129

Tabla 14. Resultados para distintos valores de N en el factor de relevancia

El algoritmo CF-I final podría quedar como sigue:

- Uso para cálculo de similitud del coeficiente de correlación de Pearson con uso de medias relativas a las asignaturas.
- Aplicación del parámetro de mejora para Pearson correspondiente al factor de relevancia, multiplicando la similitud obtenida por *n/30* si *n<30*, donde *n* es el número de alumnos compartidas por las asignaturas en cuestión.
- Elección de los 15 mejores vecinos para las predicciones.
- Predicciones realizadas redondeando al entero utilizando la suma media ajustada de las predicciones de los vecinos seleccionados.
- Aplicación del parámetro de mejora para la suma media ajustada correspondiente a la amplificación de casos.

Con estos parámetros, y para el total de pruebas realizadas, se ha obtenido un error absoluto medio de 1,0003 (9,09%) y una cobertura del 99,7652% de las predicciones.

4.4.3 Comparativa entre CF-U y CF-I

Vamos a realizar pruebas para los mejores parámetros observados con anterioridad, comprobando el comportamiento de los algoritmos en función del tamaño del conjunto de prueba.

Comparativa entre CF-I y CF-U				
% prueba	MAE CF-I	MAE CF-U	Cobertura CF-I	Cobertura CF-U
10	0,996	0,9114	99,8224	99,5775
20	0,9997	0,9134	99,741	99,4398
30	1,007	0,9243	99,7107	99,4516

Tabla 15. Comparativa entre CF-I y CF-U

En todos los casos, podemos comprobar que los mejores resultados se obtienen para CF-U. También se puede observar que la diferencia entre el MAE obtenido por CF-I no es muy grande para los distintos conjuntos de prueba, y sin embargo en CF-U vemos que para una cantidad del 30% de asignaturas en el conjunto de prueba el sistema presenta una desmejora de en torno al 0,01, aunque la diferencia tampoco es que sea significativa.

Las coberturas obtenidas son muy similares, aunque ligeramente superiores en CF-I.

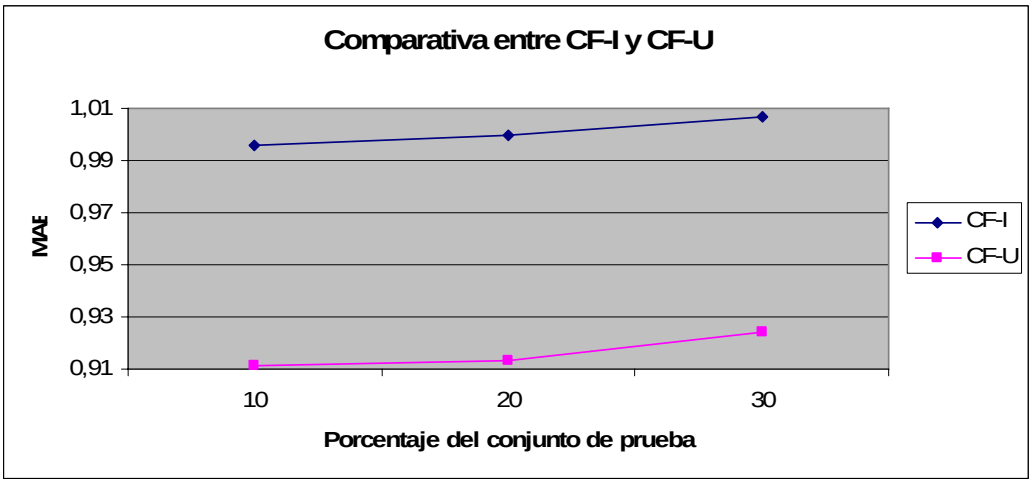


Figura 27. Gráfica comparativa entre CF-I y CF-U

Otro asunto a tener en cuenta y que puede resultar muy interesante pasa por comprobar con qué tipo de asignaturas se comportan mejor los algoritmos, si con las comunes, con las de modalidad, o con las optativas, para ver si existen diferencias significativas entre las predicciones para unas y otras asignaturas en ambas opciones.

Vamos a realizar pruebas con los parámetros ya fijados, pero exigiendo obtener predicciones sólo para un tipo determinado de asignaturas. Para ello se exigirá a los conjuntos de prueba que utilicen asignaturas únicamente de dos de los tres tipos a la hora de calcular la similitud, de forma que las predicciones se realicen independientemente para cada uno de los tipos de asignaturas que se nos presentan.

Resultados según el tipo de asignatura			
Tipo de CF	**Asignaturas**	**MAE**	**Cobertura**
CF-I	Optativas	1,1668	99,4826
CF-I	De modalidad	0,9442	99,9476
CF-I	Comunes	0,9189	99,837
CF-U	Optativas	0,8617	98,8774
CF-U	De modalidad	0,9864	99,5714
CF-U	Comunes	0,9525	99,7139

Tabla 16. Resultados según el tipo de asignatura

Es muy interesante el hecho de que las diferencias sean tan acuciadas como podemos ver en la tabla. Para el caso de CF-I, las predicciones que peor se calculan son aquellas referidas a las asignaturas optativas, para las que el error absoluto medio se dispara en aproximadamente 0,16 con respecto a la media.

En el caso de CF-U, el resultado es contrario, las mejores predicciones se realizan para estas asignaturas optativas, rebajándose la media del MAE en aproximadamente 0,05 puntos, habiendo una diferencia entre el error medio producido por CF-U y el generado por CF-I para el cálculo de predicciones en asignaturas optativas de más de 0,2 puntos.

Sin embargo, y aunque los resultados obtenidos por CF-U y CF-I sean bastante parecidos, CF-I consigue mejores predicciones para las asignaturas comunes y las asignaturas de modalidad, dándose el mejor resultado para las comunes.

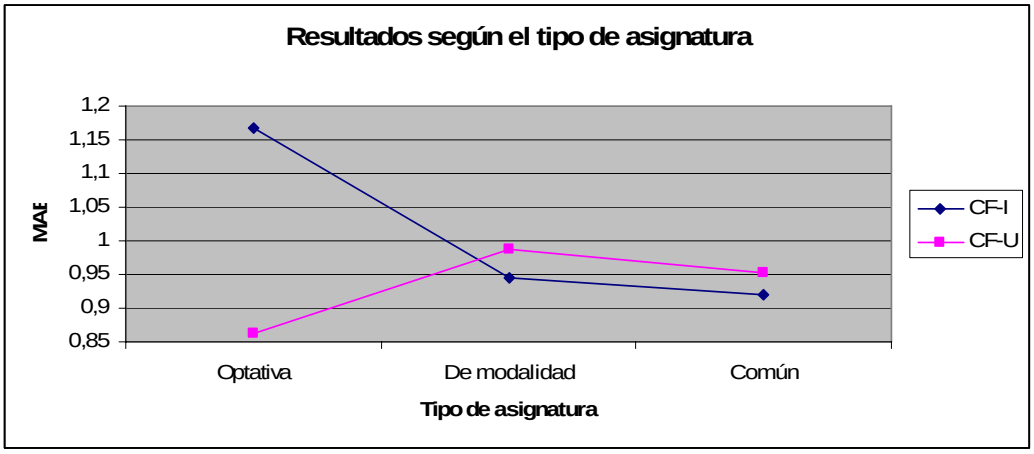


Figura 28. Resultados según el tipo de asignatura

Cabe destacar que de todas las pruebas realizadas, en predicciones puntuales el mejor resultado obtenido una vez optimizados los parámetros ha sido para asignaturas optativas mediante CF-U, obteniéndose un MAE de 0,6337 para asignaturas optativas y de 0,8026 en general, con coberturas del 98,4579% y del 100% respectivamente. Para CF-I, se ha llegado a conseguir un MAE de 0,8252 para asignaturas comunes, mientras que el mejor resultado general obtenido ha sido de 0,9266, con coberturas del 99,8182% y del 99,7531% respectivamente.

Los peores resultados han sido de un MAE 1,0494 para asignaturas de modalidad mediante CF-U, con cobertura del 99,8852%, un MAE de 0,9784 con cobertura del 98,7572% para predicciones generales con CF-U, y en el caso de CF-I un MAE de 1,2936 con el 99,3921% de cobertura para asignaturas optativas, y un MAE de 1,0284 con el 99,7594% de cobertura para predicciones generales.

4.5 Otras pruebas

Se han realizado pruebas incluyendo calificaciones de 4º de ESO, aunque sólo había disponibles datos para alrededor de unos 400 alumnos más, y se ha comprobado que existe cierto peligro si tratamos de incluir datos desequilibrando

las promociones, si el porcentaje de alumnos que se incluye no es lo suficientemente grande.

Al añadir una única promoción de alumnos, intentando ver si esto aportaba mayor información a la hora de buscar la afinidad entre usuarios, los resultados obtenidos mostraron un aumento significativo del error medio y una disminución en la cobertura, hechos que pueden deberse a diversas causas:

a) estos alumnos obtienen una mayor puntuación en los cálculos de similitud al tener mayor número de asignaturas compartidas, por lo que se presuponen más afines, pero perturban la exactitud del sistema;

b) se restringe la variedad de alumnos al elegir como afines a los de la misma promoción, por lo que la probabilidad de cursar las mismas asignaturas disminuye e incluso con valores altos de *K* existen muchas previsiones desiertas.

Por ello se añadieron todos los alumnos disponibles, y se comprobó que aunque el MAE disminuía y aumentaba la cobertura, aún no alcanzaba los buenos resultados obtenidos en los anteriores experimentos. Por ello se analizaron los datos y se descubrieron los siguientes hechos, que no sólo se restringen a 4º de E.S.O., sino que también se dan en Bachillerato:

- Los datos de 4º de E.S.O. incluían asignaturas pendientes de otros cursos, que cuando eran compartidas por 2 alumnos, aumentaba la similitud mutua por el simple hecho de compartirlas.

- Existen alumnos que sólo contienen calificaciones para un curso; por ejemplo, había alumnos con evaluaciones sólo para 4º de E.S.O., dándose el caso de que el alumno decidiera no cursar el Bachillerato, o sólo para 1º de Bachillerato, si el alumno provenía de otro instituto y después de probar terminó dejando los estudios. Estos datos en su mayoría no resultan relevantes, puesto que no aportan información a la hora de realizar predicciones, y en caso de ser elegidos como vecinos, disminuirían la calidad de la predicción.

Por ello, se ha elaborado una nueva base de datos utilizando alumnos con calificaciones para 4º de E.S.O., 1º de Bachillerato y 2º de Bachillerato. Este nuevo conjunto de datos tenía un total de 1300 alumnos repartidos en 141 grupos, con 175 materias y un total de 19981 calificaciones.

Sobre este conjunto de datos se han aplicado ciertas operaciones optimización de los datos para conseguir que fueran los más relevantes posibles:

- Se han eliminado las calificaciones nulas (datos que reflejan que el alumno se matriculó en una asignatura pero no tenía calificación).

- Se han asignado como 0 los valores de las calificaciones que aparecían como no presentado.

- Se han eliminado los ítems sinónimos, agrupando estas asignaturas en una única materia.
- Se ha realizado una selección de alumnos relevantes, de forma que se han eliminado todos aquellos alumnos que únicamente tenían calificaciones para un curso.

Concluidas estas operaciones, el nuevo conjunto de datos pasó a tener un total de 744 alumnos repartidos en 141 grupos, 100 asignaturas y 15752 calificaciones.

Se han realizado pruebas para este nuevo conjunto de datos con el algoritmo optimizado para el dominio y siguiendo la metodología anteriormente explicada, habiéndose obtenido los siguientes resultados:

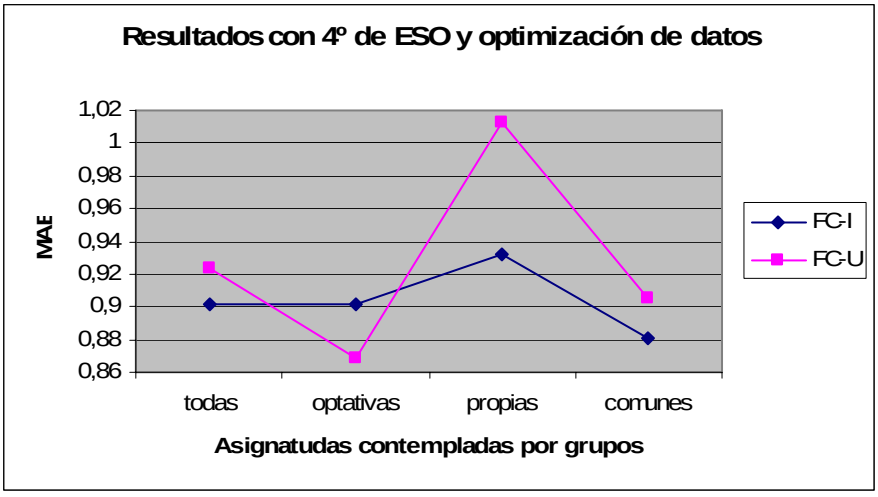


Figura 29. Resultados añadiendo 4º de ESO y optimizando datos

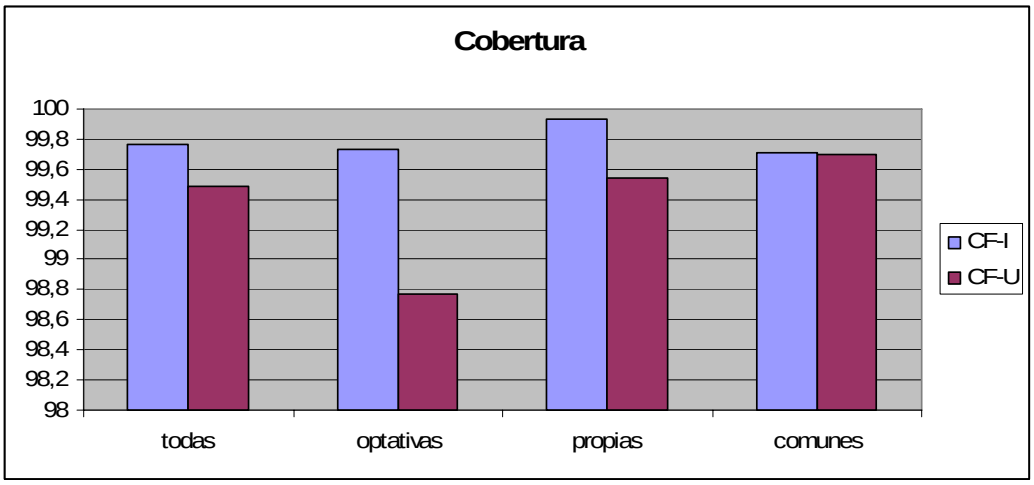


Figura 30. Cobertura añadiendo 4º de ESO y optimizando datos

Pruebas con 4º de ESO y optimización de datos				
Tipo de Algoritmo	MAE			
	todas	optativas	propias	comunes
CF-I	0,902	0,9015	0,9316	0,8803
CF-U	0,9234	0,8682	1,0125	0,9053

Tabla 17. Resultados añadiendo 4º de ESO y optimizando los datos

No debe extrañar el hecho de que el MAE general no sea la media de los errores por tipo de asignatura, puesto que el número de asignaturas de cada tipo difiere. Mencionado esto, y como podemos comprobar, el aplicar esas operaciones a los datos ha mejorado los resultados con respecto al conjunto de datos anterior, sobre todo para CF-I, y ello teniendo en cuenta el hecho de haber incluido información adicional que en teoría podría haber planteado problemas. Sin embargo, el algoritmo se ha comportado de forma adecuada, aunque en el caso de CF-U la cobertura desciende ligeramente, quizás por el hecho comentado anteriormente de haber incluido pocas promociones, y curiosamente el resultado final no mejora.

También es interesante el hecho de que se mantenga la tendencia a generar mejores recomendaciones en CF-U para las optativas, mediante CF-I para el resto.

Se está pensando en realizar un estudio más a fondo a cerca de la inclusión de datos de cursos más lejanos, para analizar el impacto de estos alumnos en el cálculo de similitudes y de nuevas predicciones, y si merece la pena más incluir sólo los datos del curso anterior, o una trayectoria más general del alumnado.

Sin embargo, para que esto sea posible se hace necesario un volumen de datos más amplio que hasta el momento no está disponible.

4.6 Adecuación de los experimentos

Como dijimos, inicialmente se han utilizado conjuntos de prueba de en torno a 800 usuarios con una cantidad de valoraciones que ronda las 13,500 para un total de 74 asignaturas. Tras añadir los alumnos de 4º de ESO disponibles y aplicar la optimización del conjunto de datos, los números no son muy distintos a los anteriores.

Así, si comparamos este volumen de datos con el que maneja cualquier sistema de recomendación comercial podríamos pensar que tal cantidad de datos no es suficiente, sin embargo debemos considerar el hecho de que contamos con un factor que juega a nuestro favor: un número limitado y muy bajo de ítems.

Al sólo tener que realizar predicciones para un máximo de 100 asignaturas la cantidad de datos requerida disminuye considerablemente, y a la vista están los resultados obtenidos.

Con respecto a las procedencias de los datos, es verdad que los datos muestran debilidad en las siguientes características:

- **Representatividad**: al estar recogidos de únicamente dos Institutos de Educación Secundaria no podemos decir que los datos sean completamente representativos si quisiéramos utilizarlos para realizar predicciones en otros centros. Sin embargo, hemos visto que los resultados son buenos, lo que nos puede inducir a que cada centro podría manejarse únicamente con sus propios datos. No quiero de todas formas aventurarme a lanzar este tipo de conclusiones sin un estudio más detenido y la evaluación de nuevas mejoras al proceso, mejoras que se comentarán en *Trabajo Futuro*.
- **Carácter cualitativo**: durante la realización de predicciones, el sistema únicamente se basa en datos cuantitativos, hecho que no debería bastar para que el sistema pueda hacer recomendaciones de fiar, puesto que como mínimo habría que tener en cuenta los gustos del alumnado, y en base a ellos decidir qué tipo de asignaturas conviene más.
- **Generalidad**: los resultados obtenidos se ciñen a predicciones para Bachillerato, valiéndonos de datos tanto de Bachillerato como de 4º de ESO. Habría que evaluar el sistema incluyendo un mayor número de cursos, e incluso podría plantearse el realizar pruebas en enseñanzas universitarias.

A pesar de estos hechos, es innegable que podemos ser optimistas frente al desarrollo de un sistema que persiga los objetivos que inicialmente nos planteamos, y que además salve estas taras de forma holgada.

4.7 Explicación de los resultados

¿Por qué, a pesar de las características inherentes a este complicado dominio, un algoritmo de tipo CF es capaz de realizar predicciones acertadas?

En realidad, aunque la idea inicial de los algoritmos CF era la de aconsejar a usuarios en base a las decisiones que tomaron usuarios con gustos parecidos al usuario activo, es decir, se trataba de recomendar en base al gusto personal de los usuarios, las buenas o malas calificaciones en gran parte de las asignaturas se puede decir que también se debe a ese ‘gusto’. Existen alumnos que están ciertamente dotados para diversas materias, y eso hace que saquen buenas calificaciones a la vez que los motiva a la hora de estudiar asignaturas de ese tipo, por lo que generalmente obtendrán unas calificaciones parecidas en ese rango de materias. Del mismo modo ocurre con aquellas que no son del gusto del

alumno, le cuesta más estudiar, no encuentra motivación, y suele obtener peores resultados.

También existen profesores que hacen de sus materias interesantes fuentes de aprendizaje, lo que estimula a los alumnos y produce que distintos alumnos suelan obtener parecidos resultados en esas asignaturas.

Por otro lado, aunque resulte arriesgado decirlo, siempre existen *estereotipos* de alumnos, como aquellos que están muy interesados en los estudios, o los que prefieren dedicar su tiempo a otras actividades más ociosas. Los algoritmos de CF son capaces de detectar tales grupos de alumnos parecidos y basarse en ellos para predecir comportamientos.

Tras analizar los resultados se ha comprobado que existen materias que siempre, en casi la totalidad de los casos, presentan tanto calificaciones como predicciones máximas (de 9 o 10). No creo que nadie se sorprenda si oye hablar del término '*María*' referido a una asignatura. Ciertamente, y sin entrar en juicios de valores, existen materias que el alumnado supera con amplia facilidad y otras que a la amplia mayoría les cuesta horrores aprobar. ¿Puede este tipo de asignaturas interferir en las predicciones? Pues más que en las predicciones, podría intervenir en las recomendaciones, sobre todo en el caso de las que siempre presentan calificaciones muy altas, puesto que siempre resultarían recomendadas en caso de ser materias optativas. En [63] se explica muy bien el término *robustez* para los algoritmos de filtrado colaborativo y cómo usuarios mal intencionados pueden actuar para hacer que ciertos ítems siempre resulten recomendados, o bien otros nunca lo sean. Sería interesante estudiar el impacto que tienen las medidas que se proponen para evitar estas circunstancias, después de haber evaluar el grado de importancia que tendrían las consecuencias derivadas de estas votaciones, en nuestro caso, de las calificaciones en estas materias.

Otro hecho curioso y que creo importante destacar es el distinto comportamiento de los algoritmos CF-U y los CF-I frente a los diversos tipos de asignaturas. Si nos fijamos en cómo funciona el CF-U, veremos que busca similitudes en alumnos a la hora de intentar estimar el resultado en nuevas materias. Es de suponer que aquellos alumnos que han tenido expedientes similares y que eligen asignaturas optativas similares, obtienen calificaciones parecidas en tales optativas; por eso el CF-U es capaz de comportarse mejor con estas asignaturas. Sin embargo, el CF-I trata de buscar similitudes entre las asignaturas que un alumno ha cursado, en base a las calificaciones que dicho alumno ha ido obteniendo. Como se dice en [12], el CF también se puede utilizar para buscar parecidos entre los ítems valorados por los distintos usuarios. De esta forma, buscará asignaturas parecidas, relacionadas, que se comporten de forma similar, y trata de estimar en base a este parecido los resultados de los alumnos. Siempre se ha dicho '*a este alumno se le dan bien los idiomas*' o '*esta chica es*

de ciencias', y eso hace referencia a este hecho, existen algunas materias que suelen dársele mejor a cada individuo. CF-I busca esta relación y la explota.

4.8 Algoritmo Colaborativo Propuesto

Tras realizar diversas pruebas de recomendaciones manuales para alumnos concretos, resultaban extraño el hecho de que aunque para CF-I el error medio era el esperado, para CF-U no bajaba de 1 al recomendar a alumnos de 4º de ESO, por lo que se decidió realizar nuevas pruebas para comprobar hasta qué punto esto era cierto y por qué.

Para ello se diseñó un nuevo sistema de pruebas de forma que dichas pruebas fueran lo más cercanas posible al funcionamiento real del sistema. El proceso de creación de los conjuntos prueba y entrenamiento varió necesariamente, ya que el sistema toma los datos de otra forma, y lo hace cogiendo a un alumno de un curso inferior, utilizando sólo los datos que se tienen de ese curso (4º de ESO, un único curso), y calculando la similitud en base únicamente a esas materias. Hablando en términos de las pruebas realizadas en el punto 4, donde el conjunto de prueba oscilaba entre el 10% y el 30% de las asignaturas, teniendo disponible entre un 70% y un 90% para los cálculos de las predicciones, ahora se obligaba al sistema a predecir calificaciones para un conjunto de prueba formado por un 70% del total aproximadamente, basándose para calcular la similitud en un conjunto de entrenamiento del 30% de materias, solo las de 4º de ESO.

Se diseñó otra metodología de pruebas más parecida al funcionamiento real en la que se eliminaban del conjunto de entrenamiento absolutamente todas las materias relativas a Bachillerato, utilizando sólo las de 4º de ESO para calcular la similitud. De esta forma, se apartaba una cierta cantidad de alumnos aleatorios de 4º de ESO para que en base a sus calificaciones en ese curso se predijeran las de los cursos siguientes

Sin mucha sorpresa, tras los temores aparecidos después de las hipótesis que se estaban barajando, se obtuvo un MAE global de 1,0675 para CF-U, casi 0,2 puntos más alto que el último obtenido en las pruebas, y con una cobertura del 98%. CF-I se mantenía en las mismas cifras prácticamente, con una cobertura del 100%.

El error en CF-U es explicable por el distinto diseño de las pruebas, pero ¿cómo el hecho de que CF-I se mantenga más o menos en los mismos valores? Fácilmente, CF-U se vio enormemente impactado por la nueva y no esperada *dispersión* de los datos, mientras que a CF-I no le afectaba para nada tal dispersión, sobre todo teniendo en cuenta que para las predicciones únicamente utiliza 15 materias, mientras que CF-U necesitaba hasta 30 alumnos similares.

En el apartado de exactitud en las predicciones por tipología de las materias se mantenía el patrón. CF-U funcionaba igualmente bien para las

optativas, y CF-I para el resto, aunque ahora CF-I y CF-U no estaban muy separados en la predicción de optativas (un 0.05 de diferencia).

Se decidió entonces poner en práctica la idea antes adelantada de un algoritmo híbrido simple de combinación de las predicciones de los dos enfoques en función del tipo de asignatura, y se barajaron diversas hipótesis.

Tras estudiar detenidamente las predicciones, se vio que en algunos casos, cuando se producían errores en la predicción de las asignaturas optativas se solía cometer un error de uno o dos puntos por parte de ambos modelos indistintamente y de forma opuesta, por lo que se decidió aplicar la media para paliar este efecto en este tipo de asignaturas. Partiendo de esta idea, se pensó en otro enfoque que utilizaría la media de CF-U y CF-I únicamente cuando la diferencia entre ambas predicciones fuera mayor que 2.

Por otro lado, se ha tratado de estimar un factor de confianza para las predicciones. Este factor de confianza se ha calculado partiendo de la idea de que cuantos más vecinos se utilizaran en una predicción, más fiable sería esta predicción. De este modo, para cada predicción se podría escoger aquella cuyo factor de confianza fuera más elevado.

También se ha probado el enfoque basado en utilizar CF-U para optativas, CF-I para propias, y la media de ambos para las comunes, método que se apuntaba en el punto 4, o utilizar CF-I para las de modalidad y CF-U para el resto, diversas combinaciones de sumas ponderadas.

Sin embargo, las pruebas realizadas con estas nuevas propuestas de algoritmos no dieron los resultados esperados, y seguía comportándose igual y en algunos casos mejor CF-I, por lo que será esta técnica, en su versión optimizada vista anteriormente, la que se implementará en un sistema de prueba.

5. DISCUSIÓN SOBRE LAS PRUEBAS EXPERIMENTALES

5 DISCUSIÓN SOBRE LAS PRUEBAS EXPERIMENTALES

Hemos estudiado las distintas técnicas usadas en los sistemas de recomendación que se basan en filtrado colaborativo, para posteriormente tratar de aplicar tales algoritmos a un dominio particular en el que lo más característico es la forma en la que las valoraciones se realizan, puesto que se trata de valoraciones que no son asignadas directamente por el usuario (explícitas) ni obtenidas automáticamente por el sistema (implícitas), sino determinadas por terceras personas en función del comportamiento, desempeño y resultados obtenidos por un alumno a lo largo del curso para una asignatura.

Se han realizado diversas pruebas aplicando parámetros de mejora y usado tanto filtrado colaborativo basado en usuario como en ítem, y se han mostrado los resultados obtenidos.

Ahora se hace necesario analizar hasta qué punto estos resultados pueden sernos útiles a la hora de construir un sistema de recomendación que oriente al alumnado a crear su propio itinerario educativo.

Para el caso de CF-U se ha comprobado que la configuración que mejor funciona es aquella que usa el coeficiente de correlación de Pearson con media fija como medida de similitud aplicando el factor de relevancia (para un máximo de 50 materias comunes), y usando 30 vecinos para realizar las predicciones, que se calculan mediante suma media ajustada con amplificación de casos. El error medio ronda 0.92 puntos en una escala de 11 valores (de 0 a 10) y soporte 10, y la cobertura no desciende del 98 o 99%.

En CF-I, la aproximación que mejor ha funcionado ha sido usando el coeficiente de correlación de Pearson puro (sin fijar la media) aplicando el factor de relevancia (para un máximo de 30 alumnos comunes) y usando 15 materias vecinas para la realización de predicciones, calculadas como en el caso anterior mediante la suma media ajustada con amplificación de casos. En este caso aunque la cobertura no baja del 99%, el error medio oscila alrededor de 0,895 (8,1%).

Hemos visto que tras optimizar los datos ambos algoritmos se comportan de forma similar siendo más exacto CF-I, aunque CF-U funciona mucho mejor con las materias optativas que con las propias, y CF-I se mantiene más preciso para las materias comunes y las de modalidad; cabría pensar que se puede realizar un algoritmo híbrido CF-UI que tomara del CF-U las calificaciones para las optativas, y del CF-I las propias de modalidad, y una combinación de ambas predicciones para las materias comunes, sin embargo, se han realizado diversas pruebas y no se mejoran significativamente los resultados con respecto a CF-I.

En definitiva, si nuestro objetivo era evaluar la viabilidad del uso de algoritmos colaborativos en este dominio concreto, a lo largo de la exposición de los resultados se ha demostrado que un sistema basado en filtrado colaborativo

sería capaz de predecir con una exactitud aceptable para qué asignaturas un alumno podría estar más capacitado, e incluso para cuáles necesitaría cierto refuerzo académico o le costarían más trabajo, en definitiva, sería capaz de ayudar al alumnado en general en la difícil decisión de elegir un futuro. El margen del error medio, de en torno a sólo un 8 o un 9%, nos permite pensar en la posibilidad de crear tal sistema.

Alguien podría pensar que el hecho de presentar las predicciones no sería ni de ayuda a profesores o alumnos, ni tan siquiera adecuado. Cierto, ese no sería el tipo de orientación que se pretende dar. Sin embargo, estas predicciones si pueden ayudarnos evaluando para qué modalidades el alumno estaría mejor capacitado, dentro de la modalidad qué materias se le darían mejor, y también qué materias optativas serían recomendables. Está claro que basándonos en predicciones, podríamos decir que sería como decirle al alumno '*elige estas materias porque vas a sacar mejores notas*'. En parte sí, pero no hay que perder de vista que los algoritmos colaborativos buscan también similitudes en ítems escogidos, lo que quiere decir que si un alumno elige un perfil de tecnología, difícilmente se le recomendará una optativa relacionada digamos con la literatura, puesto que en el pasado **los alumnos con un perfil parecido no la eligieron**, y por tanto el sistema no será capaz de recomendarla.

Si nos fijamos en el comportamiento de otros sistemas comerciales es fácil de entender esto; por ejemplo, Filmaffinity jamás recomendará a un usuario una película que sus almas gemelas no hayan votado. Este hecho resulta interesante en nuestro dominio porque define en cierta medida el itinerario o perfil que un alumno pueda tomar.

Para terminar, ¿qué podríamos recomendar con estos algoritmos y cómo trabajaríamos con las predicciones para hacerlo? Veámoslo por partes:

- **Recomendación de modalidad**: el sistema podría mostrar al alumno las 2 modalidades que mejor se ajusten a su perfil. Para ello, en base a las predicciones para las asignaturas de modalidad se podría calcular la media de puntuación para cada una de las modalidades y elaborar una lista ordenada con dichas modalidades.
- **Recomendación de materias propias**: el sistema mostraría una lista ordenada por preferencia de las materias propias de cada modalidad, basándose en las predicciones realizadas.
- **Recomendación de materias optativas**: igual que en el caso anterior, se ofrecería una lista ordenada con las materias optativas más recomendadas.
- **Consejo para reforzar el estudio en materias comunes**: puesto que disponemos también de predicciones para las materias comunes, podemos aprovechar tales predicciones para avisar sobre aquellas

asignaturas que el alumno está obligado a cursar y en las que debería esforzarse más porque pudiera encontrar dificultades.

¿Cómo tener en cuenta el hecho de que puede haber asignaturas que siempre presenten un nivel de recomendación más alto que el resto? Se puede lograr estableciendo grados de recomendación para las materias, y mostrando los nombres de las que superan un umbral, junto con el grado de recomendación. Este grado de recomendación puede ser un porcentaje (de 0 a 100) o una etiqueta lingüísticas (Muy recomendada, recomendada, poco recomendada, etc.).

Ni que decir tiene que las sugerencias presentadas más arriba son ampliamente mejorables utilizando más información de tipo cualitativo y proponiendo un sistema que *navegue* por los posibles itinerarios que el alumno pueda elegir, aconsejando y orientando a su vez sobre las elecciones más virtualmente adecuadas.

6. ORIEB. IMPLEMENTACIÓN DE UN SISTEMA DE ORIENTACIÓN PARA EL BACHILLERATO

6 ORIEB. IMPLEMENTACIÓN DE UN SISTEMA DE ORIENTACIÓN PARA EL BACHILLERATO

Para tener una idea básica pero clara de cómo podría plantearse un sistema de recomendación basado en algoritmos colaborativos, que se aprovechara de los resultados positivos que hemos obtenido tras el análisis de las pruebas realizadas, se ha implementado un pequeño sistema de prueba, OrieB (Web de Orientación para el Bachillerato), en el que se permite a un usuario solicitar orientación sobre la modalidad y/o materias a cursar en un año determinado.

Empezaremos este apartado haciendo una presentación general del sistema, mostrando las distintas opciones que posee y familiarizándonos con la interfaz, para posteriormente entrar a explicar la forma en la que se muestran las recomendaciones, cómo estas recomendaciones se obtienen, el grado de confianza que aportan dichas recomendaciones, etc.

6.1 Presentación del sistema OrieB

OrieB es un sistema web de recomendación basado en filtrado colaborativo, implementado sobre un servidor Apache, con tecnología ADODB [64] basada en PHP 5 [65] y MySQL 5 [66] como base de datos soporte para la información necesaria . Para la implementación HTML se hace uso de los estándares HTML 4.01 y XHTML [67], así como de Hojas de Estilo en Cascada o CSS 2.1 [68, 69].

El diseño de la base de datos se basa en el Modelo Relacional. Podemos ver un diagrama Entidad-Relación orientativo en la Figura 31, en el que aparecen únicamente los atributos menos intuitivos:

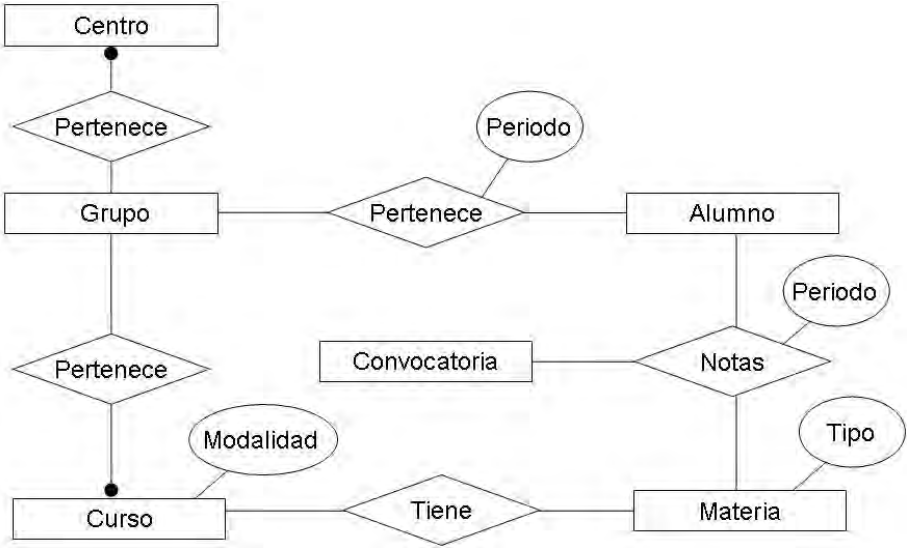


Figura 31. Diagrama Entidad-Relación orientativo

La modalidad se indica mediante el curso y el tipo de materia concreto; si es de tipo *propia de modalidad*, se buscan los cursos en los que aparece determinándose de este modo las modalidades en las que la materia se oferta.

La totalidad de los cálculos se realizan sobre el propio sistema de bases de datos en el SQL particular de MySQL, por lo que el servidor web solo tiene que solicitar los datos necesarios y presentar los resultados requeridos.

Los datos utilizados finalmente para esta implementación, después de haber sido optimizados como se comentó en el punto *4.5*, corresponden a un total de 744 alumnos almacenados de forma anónima y provenientes de dos centros educativos distintos. Se han contemplado alumnos de 2 convocatorias para 4º de ESO, y 10 convocatorias para Bachillerato, con un total de 141 grupos. El número total de materias es de 100 repartidas entre los tres cursos, formadas por 27 materias comunes, 32 propias de modalidad, y 41 asignaturas optativas. El número total de calificaciones es de 15752.

El algoritmo colaborativo a utilizar es el explicado en el punto *4.8*, aunque con una pequeña aportación. Como dijimos en la discusión, el algoritmo basado en CF-U tiene un efecto ventajoso a la hora de realizar las predicciones, y que no tiene que ver con la exactitud de las mismas: no se obtienen predicciones para aquellas materias que alumnos con las mismas preferencias y parecidos expedientes no eligieron. Esto aporta una información de carácter orientativo al alumno, en la medida en la que se le concretan más claramente las opciones a contemplar.

Por esto se ha decidido que, aunque se usará el algoritmo CF-I propuesto, se van a eliminar explícitamente de las predicciones de éste aquellas que CF-U no fuera capaz de predecir. Es decir, la predicción se realizaría mediante CF-I, pero la cobertura la definiría CF-U.

Aunque por alguna circunstancia se piense que esto pueda repercutir en una degradación de la cobertura de CF-I, hagamos notar que para CF-U la cobertura ronda el 98-99% (Figura 30). Será ese 1-2%, que incluso si queremos podemos aumentar más simplemente disminuyendo el valor de *K* (ver Tabla 9 para los distintos valores de *K* en CF-U), el que defina el perfil más claramente.

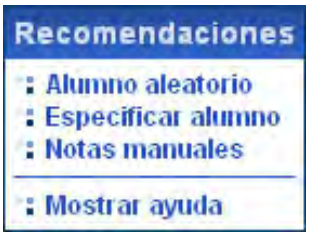
6.2 Interfaz del sitio web

Pasemos a dar una visión general del funcionamiento del sistema en sí, navegando por sus distintas opciones y obteniendo los distintos tipos de recomendaciones que ofrece.

El sitio web tiene una interfaz sencilla y amigable, en cuya página principal se da la bienvenida al usuario y se ofrece una pequeña visión del sistema y para qué puede ser útil (Figura 32).

Figura 32. Página de Bienvenida de OrieB

El menú *Recomendaciones* que aparece a la izquierda posee las siguientes opciones:



- **Alumno aleatorio**: tras elegir el curso para el que se solicita la recomendación, el sistema escogerá un alumno de forma aleatoria del curso anterior y sin tener en cuenta las posibles calificaciones que ese alumno pudiera obtener en el curso objetivo y otros posteriores, se realizan las recomendaciones pertinentes. La interfaz que solicita los datos del curso es muy intuitiva, y parecida en los 3 tipos de recomendaciones ofrecidos. Esta opción se utiliza únicamente para probar la funcionalidad del sistema.
- **Especificar alumno**: se solicita un identificador de alumno para realizar, del mismo modo que en el caso anterior pero teniendo en cuenta este alumno, las recomendaciones para el curso solicitado. Podemos ver la interfaz de solicitud de datos para esta recomendación en la Figura 33:

Figura 33. Interfaz para alumnos con identificación

- **Notas manuales**: en este caso tras indicar el curso para el que se desea la recomendación, el alumno introducirá las calificaciones que ha obtenido en el último curso. Es importante tener en cuenta que el sistema no contempla el número total de notas introducido, lo que quiere decir que funcionará aunque se le introduzca sólo un valor. Se pretende de esta forma que se puedan obtener recomendaciones más generales, como por ejemplo, para un alumno que se le de muy bien un tipo de materias, o mal otro; también es cierto que posiblemente las recomendaciones ofrecidas en estos casos sean menos exactas. Podemos ver la interfaz de introducción de notas de forma manual de las materias correspondientes a 1º de Bachillerato en la Figura 34 (para las materias de 4º de ESO la interfaz es análoga):

Figura 34. Introducción de calificaciones de forma manual

- **Mostrar ayuda**: en la ayuda aparece una explicación útil para interpretar las recomendaciones.

6.3 Explicación de las Recomendaciones

Debe quedar muy claro que éste sistema es orientativo, y que la decisión final no depende de él, sino que es una ayuda que puede permitir explorar de otro modo las distintas posibilidades que se presentan para el futuro más próximo del alumnado.

El sistema muestra, en función del curso que se solicita, diversos tipos de recomendaciones, pero antes de entrar en este tema es necesario aclarar el concepto de *explicación de las recomendaciones*.

Aunque los sistemas basados en filtrado colaborativo han demostrado ser suficientemente exactos en dominios relativos al entretenimiento, todavía no han tenido suficiente éxito en ámbitos con un alto riesgo asociado a la decisión [26]. Existen diversas razones que explican por qué no se confía en estos sistemas para este tipo de dominios: primero, estos sistemas calculan sus predicciones basándose en procesos que son aproximaciones heurísticas de los procesos humanos. Segundo, estos sistemas basan sus cálculos en unos datos dispersos, escasos y generalmente incompletos.

Aunque la segunda de las causas en nuestro ámbito no se cumple, la primera sigue estando presente, y aunque generalmente las recomendaciones generadas sean acertadas, existe la posibilidad de que contadas situaciones los consejos resulten estar completamente equivocados. Una forma de solucionar esto sería explicar al usuario cómo se ha llegado a obtener la recomendación, de manera que pudiera considerar si la conclusión es acertada entendiendo el proceso mediante el cual se ha llegado a ella, y conocer así sus puntos fuertes y sus debilidades.

En OrieB, para explicar las recomendaciones y que el usuario pueda valorar su fiabilidad se introducen 2 términos, interés y confianza de una recomendación, que pasamos a explicar a continuación.

6.3.1 Interés

En las recomendaciones existe un factor que es el grado de **interés** de una asignatura, es decir, el **grado en el que una materia sería una buena opción a contemplar o no**. Este interés viene representado mediante unas imágenes como las siguientes:

-  Interés máximo, correspondiente a tres manos con el pulgar hacia arriba, que podrá ir decreciendo si van disminuyendo la cantidad de manos con el pulgar hacia arriba que aparecen.

- 👎👎👎 Interés mínimo, representado como tres manos con el pulgar hacia abajo; va aumentando el interés conforme disminuye la cantidad de manos con el pulgar hacia abajo que se ven, aunque debemos tener en cuenta que seguirían teniendo un interés bajo.

6.3.2 Confianza

Por otro lado, se usa el término **confianza** para hacer referencia a la **medida en la que nos podemos fiar de una recomendación en concreto**. Esto quiere decir que si una materia aparece la primera al tener el más alto interés, pero con un grado de confianza del 5%, quizá sea mejor hacer más caso a la segunda en la lista si su confianza es del 60%, aunque tenga un interés algo menor.

Como hemos dicho, es muy importante no confundir la confianza con el interés. Para que quede claro, volvemos a decir que el interés expresa el grado en el que sería bueno cursar una asignatura u otra, y la confianza el grado en el que nos podemos fiar de ese interés.

6.3.3 Tipos de recomendaciones

Con respecto a las recomendaciones, las encontrarás divididas en varias secciones.

Primero, aunque únicamente para las recomendaciones solicitadas con respecto a 1º de Bachillerato, aparece una r**ecomendación sobre el perfil o modalidad** a escoger. En esta recomendación se muestra una lista ordenada de la modalidad más recomendada a la menos recomendada.

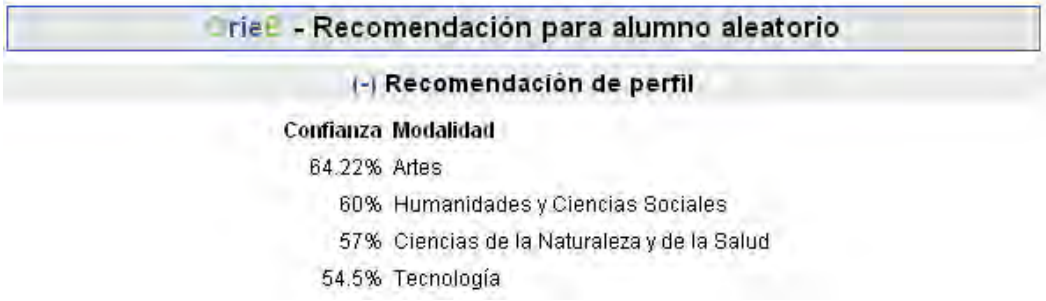


Figura 35. Recomendación de modalidad o perfil

Como podemos ver en la Figura 35, a la izquierda del título de la recomendación aparece un enlace con un signo de restar entre paréntesis (-); si hacemos clic en él los resultados se colapsarán y sólo se mostrará el título, y si ahora hacemos clic en el signo, que ha cambiado por uno de sumar (+), se restaurarán los datos. Esto se repite para todas las recomendaciones, de forma que se hace más fácil la navegación.

Posteriormente se muestran las **recomendaciones sobre materias propias de modalidad**:

Figura 36. Recomendación de asignaturas propias de modalidad

En la Figura 36 aparece, para cada una de las modalidades, una lista de materias ordenadas de la más recomendada a la menos recomendada, junto con su factor de interés y de confianza. El hecho de que aparezcan todas las modalidades cuando ya se ha recomendado una modalidad tiene sentido porque, como hemos comentado, el sistema es meramente orientador, por lo que no sería correcto limitarse a mostrar únicamente las materias aquella modalidad más recomendada de la modalidad. El alumno siempre es el que elige, el sistema únicamente le ayuda a contemplar opciones

Como podemos ver en la Figura 36 y ya explicamos más arriba, se ofrece una medida del interés de forma gráfica. En el caso de este alumno, se le había recomendado la modalidad de Artes pero también se muestran las demás, aunque como se puede ver únicamente para la modalidad de Artes el interés es ligeramente elevado. Incluso para esa modalidad existe una materia que puede resultar comprometida, Dibujo Técnico. En este caso, de decantarse el alumno por la modalidad de Artes, estaría obligado legalmente a elegir las 3 materias, sin embargo estaría prevenido de que en Dibujo Técnico podría tener problemas, y debería esforzarse para obtener un buen resultado.

Después se muestran las **recomendaciones sobre materias optativas**, que tienen la misma estructura que en el caso anterior, pero lógicamente con materias optativas.

En este caso el alumno generalmente si tiene más alternativas a elegir que el número legal de materias por el que debe optar, de forma que si quisiera podría simplemente obviar aquellas para las que el interés es peor, sin perder de vista el grado de confianza aportado por la recomendación.

En la Figura 37 podemos ver cómo la asignatura *Informática Aplicada* presenta un interés alto, y una confianza aceptable, por lo que sería una buena opción, mientras que *Francés (Segundo Idioma)* presenta un interés bajo a la vez que una confianza alta, lo que quiere decir que podría no ser una buena elección el cursar tal materia.

Con respecto a la recomendación de *Inglés (Segundo Idioma)*, el factor de confianza indica que el grado de interés bajo obtenido no es tan fiable como en el caso de *Francés*, sin embargo sigue siendo representativa la recomendación en tanto en cuanto hay otras materias con mayor interés y mayor factor de confianza.

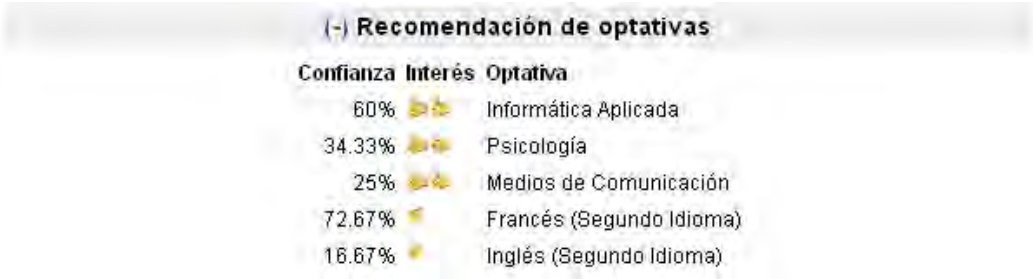


Figura 37. Recomendación de materias optativas

Para terminar con las recomendaciones, el sistema muestra aquellas **asignaturas comunes a dedicar una atención especial** (Figura 38). Puesto que como hemos estudiado, las previsiones sobre materias también pueden aportar predicciones con calificaciones suspensas, OrieB también es capaz de avisar sobre las materias en las que el alumno puede encontrar ciertos problemas, de manera que esté preparado y sobre aviso. Se indicarán aquellas asignaturas, que pueden resultar más complicadas y que con un poquito de esfuerzo extra podrán superarse sin problemas.

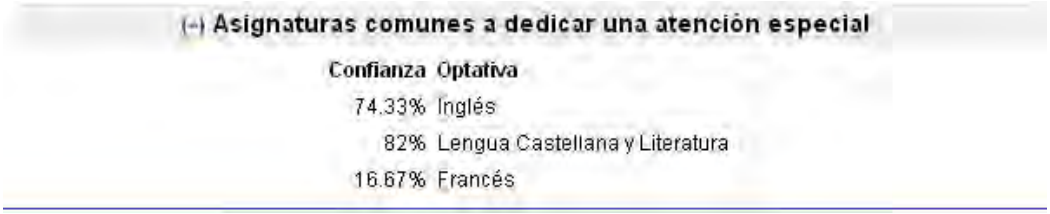


Figura 38. Materias detectadas como conflictivas o con necesidad de refuerzo

Vamos a mostrar en las siguientes imágenes un ejemplo de predicción completa para un alumno, para discutir los resultados y hablar de otro tipo de información, cuyo interés es únicamente el de comprobar la exactitud de los datos, que se muestra para algunos alumnos.

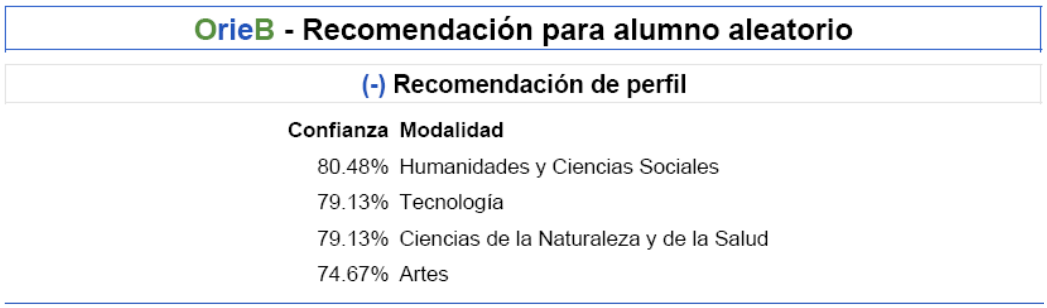


Figura 39. Ejemplo: Recomendación de perfil

La modalidad más recomendada ha resultado ser Humanidades y Ciencias Sociales, aunque seguida muy de cerca por el resto (Figura 39).

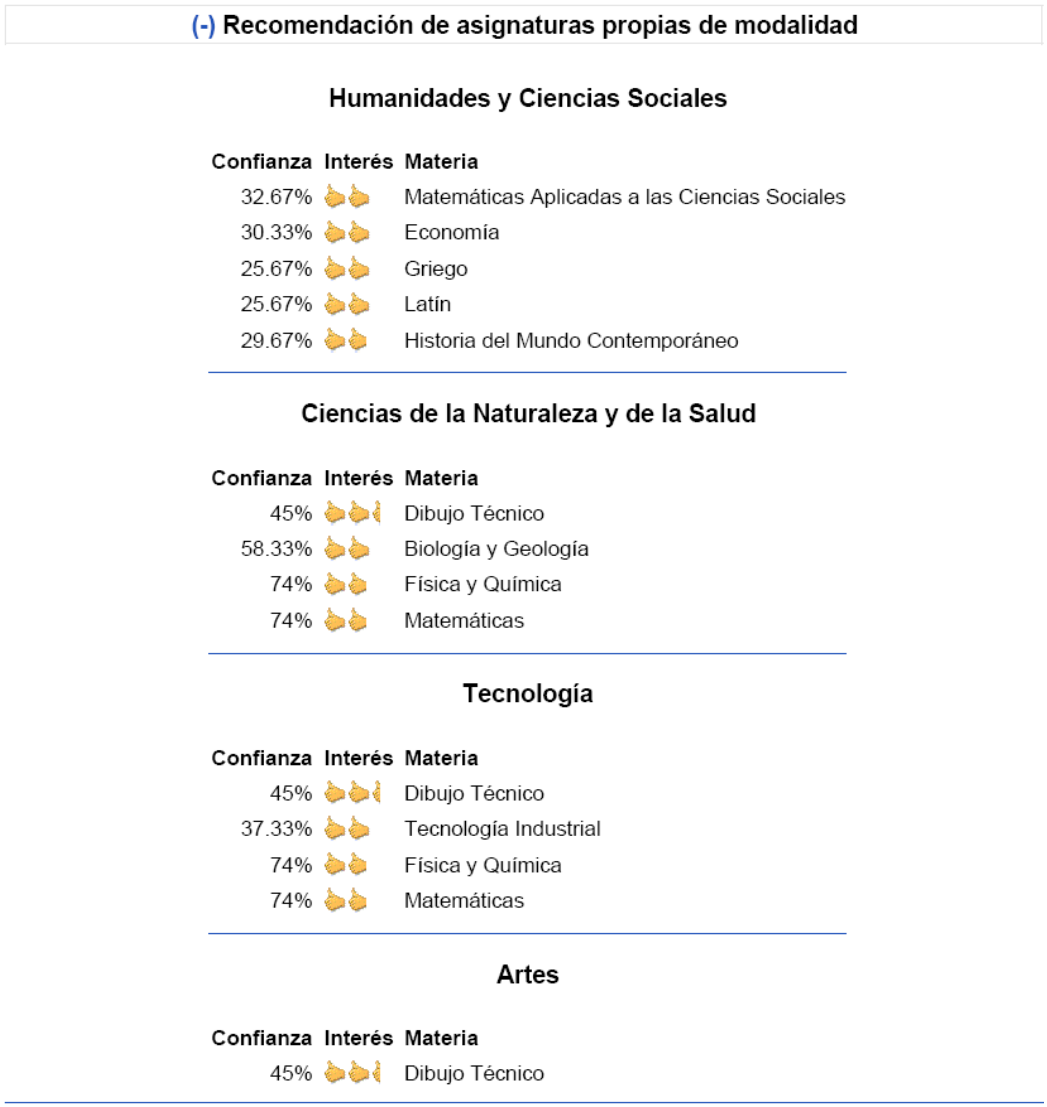


Figura 40. Ejemplo: Recomendación de materias propias de modalidad

Las materias más recomendadas para la modalidad de Humanidades fueron Matemáticas Aplicadas a las Ciencias Sociales, y Economía, mientras que la menos recomendada ha sido Historia del Mundo Contemporáneo (Figura 40).

En las otras modalidades existen materias concretas muy recomendadas (sobre todo Dibujo Técnico), pero para el resto el interés de la recomendación es inferior, y más importante, no se recomiendan todas las materias propias del perfil, es más, en Artes sólo se recomienda una asignatura, lo que quiere decir que los alumnos con expediente similar al del alumno objetivo y que en el pasado obtuvieron buenos resultados, no han escogido nunca la modalidad de Artes.

Este hecho es significativo, porque como vemos no sólo podemos obtener recomendaciones por medio de las predicciones, sino por el hecho de omisión de ciertos ítems en las predicciones. El sistema tiene este factor en cuenta a la hora de calcular la confianza y el interés como se explicará más adelante.

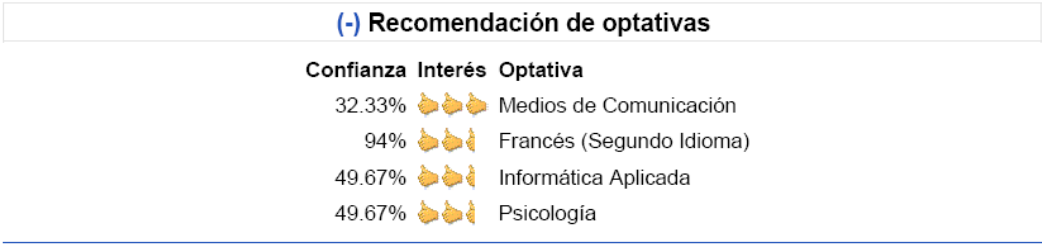


Figura 41. Ejemplo: Recomendación de optativas

La optativa más recomendada es Medios de comunicación (Figura 41), aunque con una confianza relativamente baja si la comparamos con la segunda más recomendada, Francés (Segundo Idioma), cuyo interés no es que sea bajo precisamente.

A este alumno concreto no se le previó dificultad en ninguna materia, por lo que no se mostrará esa recomendación. Sin embargo, tenemos la posibilidad de comprobar qué pasó en la realidad, debido a que se disponen de las calificaciones reales de ese alumno para los cursos de 1º y 2º de Bachillerato.

En la Figura 42 podemos ver el otro tipo de información que el sistema muestra, cuyo carácter, como antes hemos dicho, es más informativo y dirigido a evaluar la bondad y exactitud de las recomendaciones, aunque **en el sistema real estos datos no deben mostrarse**.

En la imagen podemos ver la confianza de la predicción, el curso en el que se estudia la materia, su nombre, la calificación estimada y la real. Al final también podemos ver el error medio obtenido en las previsiones, que en este caso es de 0.579 aproximadamente.

Es interesante contemplar el hecho de que este alumno realmente escogió la modalidad de Humanidades y Ciencias Sociales, también escogió las dos materias propias de la modalidad más recomendadas, Economía y Matemáticas

Aplicadas a las Ciencias Sociales, en las que obtuvo un 8 y para las que se preveía un 7, y sin embargo escogió Historia del Mundo Contemporáneo, para la que se había estimado un 6 y obtuvo un 7. Aunque la diferencia de estas predicciones con la calificación real es de 1 punto, sin embargo el grado de acierto del sistema a la hora de la recomendación es importante, puesto que el alumno obtuvo mejores resultados en aquellas materias mostradas con mayor grado de interés, y una menor calificación en las que el sistema ofrecía menor interés.

(-) Comparativa de predicciones y notas reales

Confianza	Curso	Materia	Previsión	Real
94%	1º Bach.	Francés (Segundo Idioma)	8	8
92.33%	1º Bach.	Religión y Moral Católica	9	10
91%	1º Bach.	Filosofía	7	8
91%	1º Bach.	Lengua Castellana y Literatura	7	7
88%	1º Bach.	Educación Física	6	6
49.67%	1º Bach.	Informática Aplicada	8	9
32.67%	1º Bach.	Matemáticas Aplicadas a las Ciencias Sociales	7	8
30.33%	1º Bach.	Economía	7	8
29.67%	1º Bach.	Historia del Mundo Contemporáneo	6	7
71.33%	2º BACH	Religión y Moral Católica	9	10
70.67%	2º BACH	Francés (Segundo Idioma)	8	8
67.67%	2º BACH	Historia	7	6
67.67%	2º BACH	Inglés	7	7
67.67%	2º BACH	Lengua Castellana y Literatura	7	8
67.67%	2º BACH	Filosofía	7	7
33.33%	2º BACH	Geografía	8	8
31.67%	2º BACH	Iniciación Técnico-Práctica	9	9
30.33%	2º BACH	Matemáticas Aplicadas a las Ciencias Sociales	7	6
28%	2º BACH	Economía y Organización de Empresas	7	6

Error medio de las previsiones: 0.578947368421

Figura 42. Ejemplo: Comparación de previsiones y notas reales

Las optativas elegidas se encuentran dentro de las recomendadas por el sistema, y el alumno obtuvo en ambas una calificación de 8, coincidiendo en el caso de Francés (Segundo Idioma) con la previsión del sistema.

Podemos decir que en este caso el sistema se comportó de una forma adecuada y fiable, y que incluso podría haber ayudado al alumno a considerar otras materias (Medios de Comunicación, Latín o Griego) y obtener unos conocimientos más adecuados y un mejor expediente, aunque, y eso es lo bueno del sistema, el hecho de únicamente orientar permite al alumno contemplar sus propias preferencias a la hora de elegir las materias a cursar.

6.4 Decisiones de implementación

A la hora de realizar las recomendaciones, existen tres tipos diferenciados de recomendaciones a tener en cuenta:

- La **modalidad** a cursar.
- Las **asignaturas** o materias a escoger, ya sean optativas o propias de la modalidad.
- Las previsiones de **refuerzo académico** para ciertas materias o asignaturas.

También deberemos tener en cuenta cómo tomar en consideración la posibilidad de encontrarnos *falsos negativos* y *falsos positivos*, tomando las medidas oportunas para minimizar su efecto.

6.4.1 Falsos positivos y negativos

Empezaremos con esto último, ya que las medidas a tomar para evitar estos problemas son comunes a los tres tipos de recomendaciones que se realizan.

Recordamos que en nuestro caso un falso positivo ocurriría cuando se recomienda o bien una materia o bien una modalidad que no debería haber sido recomendada. Por otro lado, un falso negativo haría referencia al hecho de no ofrecer recomendación sobre algún elemento, cuando en realidad debería haberse hecho.

Para evitar ambos problemas se han tomado las siguientes medidas:

- Cálculo de un factor de confianza para las recomendaciones: para todas las recomendaciones aportadas, se indica el grado de confianza que ofrece cada recomendación.
- El factor de confianza se calcula de forma doble: primero la confianza de la recomendación referida a cada materia, y posteriormente en lo concerniente a la modalidad. El cálculo de la confianza se explicará individualmente para cada caso cuando hablemos de la forma en la que se realizan las recomendaciones.
- Como el número de materias por curso y modalidad es finito y pequeño, se muestran al alumno todas las posibilidades de elección con el grado de interés asociado y la confianza que ofrece la recomendación. De este modo, el alumno puede considerar todas las posibilidades, aunque ofrezcan un menor interés, evaluando también la confianza de dicho interés.

De éste modo, cuando nos enfrentemos a la posibilidad de un falso positivo o de un falso negativo, tomará un papel crucial la propia decisión del

alumno que tendrá en cuenta también sus gustos, sus preferencias, y sus creencias a la hora de tomar una decisión.

6.4.2 Implementación de las recomendaciones

Como ya sabemos, para esta prueba sólo vamos a poder basarnos en la información cuantitativa que nos proporcionan las predicciones que el sistema es capaz de realizar. Por ello, tendremos que estudiar la manera en la que dichas predicciones pueden ayudar en el proceso de recomendación.

Vamos a ver por separado cómo se calculan los 3 tipos de recomendación que hemos determinado.

Recomendación de modalidad

Podemos suponer, que una modalidad concreta será tanto más adecuada para un alumno cuanto mayores sean las predicciones obtenidas para todas las asignaturas propias de dicha modalidad. Por esto mismo, uno de los factores que vamos a tener en cuenta a la hora de decidir el orden en el que se recomendarán las modalidades será la media obtenida por el alumno en las predicciones para esas materias de modalidad por curso.

Sin embargo, no sólo podemos basarnos en esto para proporcionar una recomendación. ¿Qué pasaría si un alumno obtiene una previsión de 10 en Dibujo Artístico solamente? Pues que para toda la modalidad de Artes obtendría una recomendación de 10. Sin embargo, existen 9 materias propias de dicha modalidad. ¿Cómo podemos solucionar este inconveniente? Utilizaremos otro factor que tendrá en cuenta el número de materias que se han utilizado en la obtención de la media por cada curso concreto.

Por último, necesitamos tener en cuenta otro dato que evite el siguiente problema: dos modalidades con 9 asignaturas en total y predicciones para todas las asignaturas obtienen de 7,33 ambas de media de las predicciones. Sin embargo, las calificaciones previstas para las modalidades fueron:

Modalidad 1	10	4	10	10	4	4	10	4	10
Modalidad 2	8	6	8	8	6	6	8	8	8

Tabla 18. Justificación del uso de la Varianza

En base a estas previsiones, ¿sería lógico recomendar una modalidad en la que se prevén 4 suspensos del mismo modo que otra en la que se aprueba todo y con calificaciones más que aceptables? No hace falta decir la respuesta.

Es por esto que se va a considerar también como tercer factor la varianza para la media de cada modalidad.

De esta forma, construiremos la predicción en base a estos tres factores:

- $\overline{m}$: Media de todas las predicciones obtenidas para las materias de una modalidad, teniendo en cuenta cada curso del Bachillerato.
- *a*: Número de asignaturas para las que se obtuvo predicción en la modalidad contemplada.
- var: Varianza para la media de cada modalidad.

Para obtener un número del que se puede obtener una medición intuitiva, vamos a transformar los tres factores en $\overline{M}$, *A* y *VAR* respectivamente de forma que el valor de cada uno esté normalizado, comprendido entre 0 y 1. Para ello se calcularían como sigue:

$$\overline{M} = \frac{\overline{m}}{10}, \quad VAR = \frac{\text{var}}{10} \qquad \textit{Ecuaciones 10 y 11}$$

Para el cálculo de *A*, dividiremos el número de predicciones realizadas para asignaturas propias de una modalidad cada curso por el total de materias de modalidad de ese curso. Como se puede ver en el *Anexo II*, el número de materias por curso y modalidad es:

- *Artes*: 3 asignaturas en primero y 6 en segundo
- *Ciencias de la Naturaleza y de la Salud*: 4 materias en primero y 5 en segundo
- *Humanidades y Ciencias Sociales*: 5 en primero y 7 en segundo
- *Tecnología*: 4 en primero y 6 en segundo

Una vez obtenidos estos datos, se puede construir un factor de certeza o grado de recomendación para cada modalidad multiplicando ambos 3 factores. Sin embargo, existen 2 problemas con respecto a esto. El más grave de ellos se da cuando la varianza es 0, porque anula los otros dos factores completamente. Podemos solucionar fácilmente esto multiplicando por (1 – varianza), sin embargo los resultados no son, como decirlo, muy atractivos. Al estar hablando de factores que rondan normalmente una cantidad similar a 0,6 o 0,7 en los mejores casos, jamás se obtendría una recomendación de 1. Es más, en estos casos se obtendrían valores de 0,3 que, si decimos que el valor debe estar entre 0 y 1, puede resultar un valor bajo.

Una solución es efectuar una suma ponderada de los tres factores, dando la misma importancia a cada uno. Sin embargo, se ha comprobado empíricamente que los resultados son más deseables si la suma ponderada se

realiza aplicando un 60% a la calificación media, un 20% al número de materias predichas, y un 20% a la varianza.

Recomendación de materias

Con respecto a cada materia individual, existe un factor clave que mide la relevancia de cada predicción, y es el número de alumnos que se utilizaron para obtenerla. En nuestro caso, al usar un algoritmo de filtrado colaborativo basado en K-NN (los K vecinos más cercanos), el máximo número de alumnos usados para elaborar una predicción es siempre 30, por lo que podremos calcular un factor de certeza adecuado dividiendo el número de alumnos usados para la predicción entre 30.

Con respecto al interés en cursar una u otra asignatura, se puede utilizar la predicción tal cual. Pero se ha considerado que no es adecuado mostrar números en este caso, por lo que se ha elaborado la representación gráfica indicada anteriormente. Para ello, se toma la longitud total de la imagen (40 píxeles) y se muestra un porcentaje de ella calculado en función de la calificación. Si ésta es mayor que 5, la imagen está formada por las 3 manos con los pulgares hacia arriba, si es menor que 5, con los pulgares hacia abajo, y si es 5, no hay imagen.

Cuanto más se acerque la calificación a 10 más se verán las 3 manos con los pulgares hacia arriba, cuanto más se acerque a 0 más se verán los pulgares hacia abajo. El resto es fácil de intuir.

De este modo, se muestra en la recomendación todas las materias para las que se obtuvo predicción, con el grado de interés calculado y el factor de confianza asociado.

Asignaturas con necesidad de refuerzo

Las materias se consideran comprometidas y se muestran en este apartado si son materias comunes (todos deben cursarlas) y además la predicción obtenida es de 4 o menos.

Análogamente al caso anterior, se mostrará también el grado de interés aportado y la confianza obtenida.

7. CONCLUSIONES

7 CONCLUSIONES

En este trabajo se ha realizado un análisis preliminar de los sistemas de recomendación en general como herramientas capaces de ahorrar tiempo y proporcionar ayuda a la hora de tomar decisiones de diversa índole, para continuar con un estudio más profundo del ámbito de los sistemas de recomendación que utilizan el filtrado colaborativo.

Hemos visto que los sistemas colaborativos son capaces de realizar recomendaciones de ítems a usuarios basándose en la idea de que a un usuario le gustará un ítem si a otros usuarios con gustos parecidos les gustó. Este tipo de sistemas permite realizar predicciones sobre elementos de difícil tratamiento y con características no inmediatas sin necesidad de analizarlos, simplemente utilizando las valoraciones previas que otros usuarios realizaron acerca de ellos.

Tras esto se han analizado las medidas básicas de similitud y las técnicas comunes de predicción, y se ha entrado en profundidad en aquellas mejoras de este tipo de algoritmos colaborativos, ya sean mediante la utilización de ciertos parámetros (amplificación de casos, frecuencia inversa, factor de relevancia, etc.), o mediante la reformulación del problema (selección de instancias, análisis de características, etc.).

Posteriormente se ha planteado un problema para comprobar si sería beneficioso utilizar una aproximación de estos sistemas colaborativos a la hora de realizar recomendaciones personalizadas a alumnos de Bachillerato a la hora de escoger asignaturas optativas y de prever qué asignaturas comunes presentarán unas mayores dificultades de aprendizaje o necesidades específicas de refuerzo al alumno.

Expuesto el problema, se han explicado las peculiaridades del dominio concreto y se ha observado que nos encontramos con unos datos con los que el CF no está acostumbrado a trabajar, unos datos que no provienen directamente del usuario y que tampoco son recogidos automáticamente por el sistema, sino que son aportados por expertos.

Mediante una serie de experimentos se ha demostrado la viabilidad de crear un sistema basado en filtrado colaborativo capaz de orientar al alumnado tal y como hemos visto, obteniéndose en las pruebas realizadas un error absoluto medio de en torno a 0,9 puntos, resultado que se ha considerado aceptable a la hora de considerar la posibilidad de implementar tal sistema, aunque también se han mostrado los retos que habría que vencer y se ha dado a entender que dicho sistema debería incluir información adicional de tipo cualitativo que recogiera como mínimo tanto los gustos como las aptitudes de cada individuo concreto.

Demostrada la viabilidad de este tipo de algoritmos para la predicción de calificaciones, se ha implementado OrieB, un pequeño sistema de prueba que

permite ver cómo sería grosso modo un sistema de recomendación que cumpliera el cometido que nos hemos planteado.

Por último se han propuesto caminos a seguir a la hora de proporcionar mayor información a estos algoritmos, hibridando con otros métodos de recomendación, añadiendo información difusa y/o planteando tareas multiobjetivo y votaciones multidimensionales.

8. TRABAJO FUTURO

8 TRABAJO FUTURO

A lo largo de todo el trabajo hemos explicado que el propósito del mismo era comprobar la viabilidad de la aplicación de los algoritmos de filtrado colaborativo al problema de aconsejar sobre decisiones a la hora de elegir asignaturas y de dedicar esfuerzos a las mismas. También hemos ido apuntando pautas que un sistema debería tener en cuenta a la hora de realizar recomendaciones serias y de fiar, no sólo basándose en información cuantitativa, sino también en cualitativa.

Vamos a ver una lista de tareas que quedarían pendientes de pretender llevarse a cabo tal tarea:

- **Inclusión de información difusa**: Hasta el momento hemos utilizado información numérica para realizar todos los procesos requeridos por el sistema de recomendación colaborativo; sin embargo es más común, adecuado y flexible en el ámbito de las calificaciones correspondientes al entorno educativo el modelado de las mismas mediante valores cualitativos. En el futuro nos proponemos estudiar dicho modelado cualitativo mediante el uso de la Aproximación Lingüística Difusa y dado que las valoraciones utilizadas en las calificaciones no son equidistantes si nos referimos a las etiquetas lingüísticas a las que suelen ir asociadas (Suspenso, Aprobado, Bien, Notable y Sobresaliente), también intentaremos modelarlas mediante Información Lingüística No Balanceada [6] para así obtener mejores resultados.
- **Privacidad del sistema**. Puesto que se trabaja con datos sobre alumnos, es interesante utilizar el mayor número de medidas que garanticen la privacidad de esos datos. Este tema está ampliamente comentado en [70-72].
- **Recomendaciones en base a perfiles del alumnado.** El sistema, a la hora de proponer una recomendación, debería tener en cuenta diversos tipos de información relativos al usuario en cuestión, como podrían ser:
    - Gustos y preferencias del alumno.
    - Situación socio-económica del mismo.
    - Entorno socio-educativo en el que se encuadra el centro.
    - Aptitudes y actitudes del alumno.
    - Orientación profesional o posibles estudios que el alumno se plantea para el futuro.
- **Recomendaciones en base a las características de las materias**. No sólo habría que tener en cuenta al propio alumno, sino también las particularidades de cada materia, pudiendo valorarse:

  - Las distintas aptitudes que requiere una asignatura en cuestión.
  - La rama o perfil a la que pertenece la asignatura, teniendo en cuenta la orientación profesional/educativa del alumno.
  - Beneficios que aporta una asignatura en términos de las aptitudes que desarrolla. Teniendo un perfil de las aptitudes que presenta un alumno y las que desarrollan las distintas asignaturas, podemos elegir unas u otras en función de lo que se pretende desarrollar en el alumno o en los que necesita mejorar en función de su orientación hacia el futuro.

- **Hibridación de sistemas de recomendación**: Para llevar a cabo estas tareas sería necesario introducir en el sistema otras técnicas de recomendación, como aquellas basadas en contenido, en utilidad y/o las demográficas, y formas de representación que permitan modelar perfiles tanto para alumnos como para asignaturas, tales como itinerarios académicos. Con el uso de tales técnicas solventaríamos ciertos problemas como el de la dispersión, el de el nuevo-ítem, o el de tener escasa información *offline* del usuario, como se apunta en [5].
- **Recomendaciones de mayor alcance**: Con todo esto, el sistema no solo debería pensarse para recomendar perfiles, asignaturas optativas o para informar sobre qué materias pueden resultar conflictivas en bachillerato, sino también a la hora de orientar sobre perfiles académicos de una forma más general incluyendo ámbitos universitarios, caminos curriculares más complejos, salidas profesionales de los mismos, etc., de forma que partiendo de una información básica, el alumno se encontraría con un sistema guiado que le haría navegar por distintos caminos o perfiles académicos que podría tomar a la hora de elegir su futuro, a la manera de los sistemas guiados presentados en [3].
- **Uso de información *online***: En caso de plantearse un sistema así, no sólo se requeriría información *offline*, sino que se haría necesaria la interpretación de información *online* y explícita de manos del usuario. Además, las decisiones a tomar podrían plantearse para satisfacer más de un criterio, como el grado en el que se asemejan al gusto del usuario, su conveniencia conforme a sus aptitudes, etc.

ANEXO I. CURSOS DE DOCTORADO

ANEXO I. CURSOS DE DOCTORADO.

El período de docencia del que suscribe fue realizado durante el año académico 2005/2006, en el programa de doctorado "Métodos y Técnicas Avanzadas de Desarrollo de Software", siendo coordinadores del mismo D. Juan Ruiz de Miras y D. Manuel Capel Muñón, dotado con la mención de calidad y de carácter interuniversitario entre las universidades de Granada y Jaén.

Se enumeran y describen brevemente los cursos realizados en dicho período:

- **Integración de la Información en la Web Semántica.** *Curso impartido por José Samos Jiménez (ugr) y Manuel Torres Gil (ugr).* Se estudió la Web Semántica, su finalidad y las expectativas de futuro, así como las tecnologías relacionadas (XML, RDF, etc.), ontologías relacionadas y su integración. Cada alumno realizó una presentación sobre un artículo relacionado con el área, que previamente hubo que estudiar y analizar.

- **Tecnologías del Lenguaje.** *Impartido por José María Guirao Miras (ugr), Ramón López Cózar (ugr) y Teresa Martín Valdivia (uja).* Se estudió el estado del arte en este tipo de tecnologías desde varios puntos de vista: procesamiento del lenguaje natural, métodos estadísticos para el PLN, y tecnologías del habla y sistemas de diálogo. Se mostró a los alumnos en las aulas de prácticas el funcionamiento de algunos sistemas predictores del lenguaje (al estilo de los teléfonos móviles), y algunas otras aplicaciones relacionadas con el tema.

- **Integración de Técnicas Hipermedia y de Visualización Gráfica en Sistemas Inteligentes.** *Impartido por los profesores Dª Lina Guadalupe García Cabrera (uja), Luis Martínez López (uja) y Antonio J. Rueda Ruiz (uja).* Los contenidos versaron sobre los siguientes temas: Introducción a los Sistemas Inteligentes (Sistemas de Agentes, Sistemas Multi-Agente, Aplicaciones en Internet), Tecnologías Hipermedia (Modelos de Adaptación, Modelos de Navegación, Modelos Adaptativos según la estructura del Conocimiento), Interfaces y Visualización Avanzada en Entornos Web (Interfaces de Usuario, Visualización Avanzada). Se realizaron presentaciones sobre distintos temas relacionados.

- **Recuperación de Información. Búsqueda de Respuestas.** *Curso impartido por Luis Alfonso Ureña López y Manuel Palomar Sanz***.** En este curso estudiamos primero unas nociones básicas sobre recuperación de información, procesamiento del lenguaje natural y

materias relacionadas, como desambigüación, procesamiento de voz y habla... Muchas de estas clases las impartieron expertos a nivel nacional en estas materias. La última parte de este curso se destinó a los sistemas de búsqueda de respuestas (*Question Answering*).

- **Sistemas Hipermedia.** *Curso impartido por María José Rodríguez Fórtiz y Francisco Luis Gutiérrez Vela***.** En este curso estudiamos una introducción a la hipermedia, con modelos de referencia, metodologías de diseño, técnicas adaptativas, evolutivas y colaborativas con su aplicación a la hipermedia, e-learning... En la parte práctica estudiamos diversos ejemplos de sistemas hipermedia y se realizaron ejercicios de análisis de diversos sistemas hipermedia, y también de diseño con la ayuda de aplicaciones de modelado tales como WebRatio o CMap Tools.
- **Diseño de Sistemas Colaborativos.** *Curso impartido por Miguel Gea y Juan José Cañas Delgado***.** En este curso vimos una introducción a CSCW, un software de trabajo colaborativo mediante el cual se realizó la comunicación durante el curso y toda la parte práctica. En la parte teórica estudiamos unos modelos cognitivos basados en grupos, metodologías de diseño de aplicaciones cooperativas, herramientas, taxonomías y plataformas de desarrollo de sistemas Groupware... El objetivo de este curso era el estudio de los Sistemas Interactivos, los mecanismos disponibles para desarrollarlos y la aplicación a nuevos paradigmas basados en la cooperación y el trabajo en grupo. Se estudiaron importantes artículos que analizaban el estado del arte desde diferentes perspectivas y repasamos algunos modelos básicos UML para el diseño de este tipo de sistemas, para finalizar realizando un informe por grupos de lo estudiado en teoría.
- **Ingeniería de la Usabilidad.** *Curso impartido por Francisco Luis Gutiérrez Vela, Maria Luisa Rodríguez Almendros y Julio Abascal González***.** En este curso estudiamos una introducción a la usabilidad, vimos distintos tipos de modelado, como el modelado de tareas utilizando casos de uso. También estudiamos métodos de inspección de la usabilidad, las normativas y la extensión de los modelos de usuario a modelos de grupo. Cada alumno realizó un estudio y exposición sobre un tema concreto relacionado con el área.

ANEXO II. E.S.O. Y BACHILLERATO

ANEXO II. E.S.O. Y BACHILLERATO

9 EDUCACIÓN SECUNDARIA OBLIGATORIA (E.S.O.)

La Educación Secundaria Obligatoria es una etapa educativa, obligatoria y gratuita, para todos los individuos en edad escolar que completa la Educación Básica y abarca cuatro cursos académicos.

Su finalidad es transmitir a todos los alumnos los elementos básicos de la cultura, formarlo para asumir sus deberes y ejercer sus derechos y prepararlo para la incorporación a la vida activa o para acceder a la formación profesional específica de grado medio o al bachillerato. La atención a la diversidad de intereses, motivaciones, y aptitudes de los alumnos constituye el objetivo fundamental de esta etapa educativa.

La superación de esta etapa educativa supone la obtención del Graduado en Educación Secundaria.

En cualquier caso, al finalizar la etapa todos los alumnos recibirán una acreditación del centro educativo en la que consten los años cursados y las calificaciones obtenidas en las distintas áreas y materias. Esta acreditación irá acompañada de una orientación sobre el futuro académico y profesional del alumno, que en ningún caso será prescriptiva y que tendrá carácter confidencial.

Los títulos académicos y profesionales serán homologados por el Estado y expedidos por las Administraciones educativas en las condiciones establecidas en la legislación estatal y en las normas de desarrollo que al efecto se dicten.

Su duración es de cuatro años académicos.

Un alumno y sus padres pueden optar, desde el momento en que aquél cumple 16 años de edad, por dar por finalizada su escolarización obligatoria en la etapa, en cuyo caso se le extenderá la correspondiente acreditación.

9.1 Plan de Estudios

A nivel general la siguiente tabla presenta las directrices generales estatales para situar e impartir las materias correspondientes en los distintos cursos de la E.S.O.

<table>
<tr><th rowspan="2">ÁREAS y MATERIAS</th><th colspan="2">Primer ciclo</th><th colspan="2">Segundo ciclo</th></tr>
<tr><th>Curso primero</th><th>Curso segundo</th><th>Curso tercero</th><th>Curso cuarto</th></tr>
<tr><td>Comunes en todos los cursos</td><td colspan="2">• Ciencias de la Naturaleza.
• Ciencias Sociales, Geografía e Historia.
• Educación Física.
• Educación Plástica y Visual.
• Lengua Castellana y Literatura. En su caso también Lengua y Literatura de su Comunidad Autónoma.
• Lengua Extranjera.
• Matemáticas.
• Música.
• Tecnología
• Religión o Actividades de estudio (a elección, voluntaria por curso completo)</td><td colspan="2">Durante este año de la etapa los alumnos elegirán dos entre las cuatro áreas siguientes:
• Ciencias de la Naturaleza.
• Educación Plástica y Visual.
• Música.
• Tecnología.</td></tr>
<tr><td rowspan="3">Optativas</td><td colspan="4">Segunda Lengua Extranjera.</td></tr>
<tr><td colspan="2">Algunas Administraciones educativas ofrecen las Medidas de Refuerzo en Lengua y Matemáticas. Estas enseñanzas se ofrecerán exclusivamente a los alumnos que hayan presentado problemas de aprendizaje o carencias importantes que pudieran comprometer el desarrollo de las capacidades básicas instrumentales.</td><td>• Iniciación Profesional
• Cultura Clásica.</td><td>• Iniciación Profesional.
• Cultura Clásica.
• Ética</td></tr>
<tr><td colspan="4">Además de las mencionadas en la parte superior, de oferta obligada por parte de los centros, el currículo comprenderá otras materias optativas, cuya presencia y opcionalidad, permita responder a los intereses y necesidades del alumnado, ampliar las posibilidades de su orientación, facilitar su transición a la vida activa y contribuir al desarrollo de las capacidades generales a las que se refieren los objetivos de la etapa. Los alumnos deberán cursar una en cada curso. Excepcionalmente dos.</td></tr>
</table>

Observaciones a esta tabla:

- En el caso de que el área de Ciencias de la Naturaleza se organice en dos materias diferentes, "Biología y Geología", y "Física y Química", ambas contarán como dos áreas a efectos de elección.
- Las Administraciones educativas podrán disponer, también, que el área de Matemáticas, que será cursada por todos los alumnos, se organice en el cuarto curso en dos variedades de diferente contenido: Matemáticas A y Matemáticas B.
- El bloque de contenidos denominado La vida moral y la reflexión ética, incluido en el currículo del cuarto curso del área de Ciencias Sociales, Geografía e Historia, se organizará en el cuarto curso de la etapa como materia específica con la denominación de "Ética". La evaluación de estas enseñanzas se verificará de forma independiente.

10 BACHILLERATO

El Bachillerato es la última etapa de la Educación Secundaria, tiene carácter voluntario y su duración es de dos cursos, normalmente entre los 16 y los 18 años.

Tiene modalidades diferentes que permiten una preparación especializada de los alumnos (con elección de distintos itinerarios dentro de cada modalidad) para su incorporación a estudios superiores o a la vida activa. Sus finalidades son:

1. Formación general, que favorezca una mayor madurez intelectual y personal, así como una mayor capacidad para adquirir una amplia gama de saberes y habilidades.
2. Preparatoria, que asegure las bases para estudios posteriores, tanto universitarios como de formación profesional.
3. Orientadora, que permita a los alumnos ir encauzando sus preferencias e intereses.

10.1 Titulación

La superación del Bachillerato implica la obtención del Título de Bachiller.

- Para obtener el Título de Bachiller será necesaria la evaluación positiva en todas las asignaturas de la modalidad cursada.
- El título de Bachiller facultará para acceder a la formación profesional de grado superior y a los estudios universitarios. En este último caso será necesaria la superación de una Prueba de Acceso (selectividad), que, junto a las calificaciones obtenidas en el Bachillerato, valorará, con carácter objetivo, la madurez académica de los alumnos y los conocimientos adquiridos en él. Asimismo facultará para acceder a grados y estudios superiores de enseñanzas artísticas, previa superación de la correspondiente prueba.

La duración de esta etapa es de tres años. Excepcionalmente tres para la modalidad de nocturno.

La permanencia en el Bachillerato en régimen escolarizado será de cuatro años, como máximo. Esto último no afecta a los alumnos que cursen el Bachillerato por otro régimen de enseñanza, de adultos o a distancia.

La forma normal de acceso al Bachillerato pasa por estar en posesión del título de Graduado en Educación Secundaria o título equivalente.

10.2 Estructura

Comprende dos años académicos, desde los dieciséis a los dieciocho años de edad de los alumnos, y se organiza en dos cursos.

10.2.1 Modalidades

Son las siguientes:

e. Artes

f. Ciencias de la Naturaleza y de la Salud

g. Humanidades y Ciencias Sociales

h. Tecnología

Las enseñanzas del **bachillerato** están estructuradas en modalidades; unas de corte más académico y otras más profesional, facilitando así que cada alumno pueda elegir su propio itinerario formativo en función de sus capacidades e intereses académicos y profesionales

10.2.2 Itinerarios

Las enseñanzas de Bachillerato permiten a los alumnos cursar estos estudios de acuerdo con sus preferencias, en virtud de la elección de una modalidad entre las cuatro previstas, una opción dentro su modalidad y unas determinadas materias optativas. Estas sucesivas elecciones configuran el itinerario personal de cada alumno. La posibilidad de elección aumenta progresivamente de primero a segundo curso, en el que se incluyen, dentro de cada modalidad, 2 ó 3 opciones posibles que guardan, en la mayor parte de los casos, una estrecha relación con determinados estudios posteriores universitarios o profesionales. La elección de determinadas materias optativas ofrece, por otra parte, bien la posibilidad, a los alumnos que lo desean, de concurrir a las pruebas de acceso a la universidad por más de una opción, o por una opción distinta de la prevista para la modalidad o itinerario cursado, bien la posibilidad de profundizar en aspectos específicos de su modalidad, o bien ampliar la formación que pueden adquirir en la modalidad elegida, cursando como optativas materias propias de otras modalidades o materias optativas de posible interés formativo para los alumnos de todas las modalidades del Bachillerato.

Por último, los alumnos pueden en determinadas condiciones cambiar de modalidad o de itinerario en segundo curso de la etapa; y, por otra parte, pueden, si es su deseo, matricularse al finalizar la etapa de Bachillerato en aquellas materias vinculadas con alguna opción de las Pruebas de Acceso a la Universidad, no relacionada con la modalidad superada, a fin de que puedan presentarse a las mencionadas pruebas por esa opción.

10.3 Materias

El Bachillerato se organiza en:

4. materias **comunes**, para todos los alumnos independientemente de la modalidad elegida. Pretenden contribuir a la formación general del alumnado
5. materias **propias** de cada modalidad, que caracterizan a cada una de las modalidades y contribuyen a que el alumno obtenga una formación específica ligada a la modalidad elegida, y
6. materias **optativas**, que amplían la posibilidad de elección.

Los alumnos deberán cursar seis materias propias de la modalidad elegida, tres en cada curso.

Las Administraciones educativas organizarán las Modalidades distribuyendo las materias correspondientes a cada una de ellas en los dos cursos que componen el Bachillerato. Asimismo, fijarán las materias optativas del Bachillerato, así como el número de ellas que los alumnos deberán superar en cada uno de los cursos del Bachillerato.

En dicha fijación, las Administraciones educativas podrán tener también en cuenta las propuestas realizadas por los Centros educativos. En todo caso, las materias de modalidad vinculadas a cada una de las vías de acceso a estudios universitarios se impartirán en el segundo curso de Bachillerato.

Los alumnos podrán elegir como materias optativas no sólo las que resulten de lo previsto en el apartado anterior, sino también cualesquiera de las materias definidas como propias de las diferentes Modalidades, de acuerdo con lo que al efecto determinen las Administraciones educativas en función de las posibilidades de organización de los Centros.

10.4 Objetivos del Bachillerato

El Bachillerato contribuirá a desarrollar en los alumnos las siguientes capacidades:

a. Dominar la lengua castellana y la lengua oficial propia de la Comunidad Autónoma.
b. Expresarse con fluidez y corrección en una lengua extranjera.
c. Analizar y valorar críticamente las realidades del mundo contemporáneo y los antecedentes y factores que influyen en él.
d. Comprender los elementos fundamentales de la investigación y del método científico.
e. Consolidar una madurez personal, social y moral que les permita actuar de forma responsable y autónoma.

f. Participar de forma solidaria en el desarrollo y mejora de su entorno social.

g. Dominar los conocimientos científicos y tecnológicos fundamentales y las habilidades básicas propias de la modalidad escogida.

h. Desarrollar la sensibilidad artística y literaria como fuente de formación y enriquecimiento cultural.

i. Utilizar la educación física y el deporte para favorecer el desarrollo personal.

10.5 Plan de Estudios

Distribución de las materias en cada uno de los dos cursos de las distintas modalidades

<table>
<tr><td rowspan="2">Comunes</td><td>1º</td><td colspan="3">Educación Física
Filosofía I
Lengua castellana, lengua oficial propia de las CC.AA. y Literatura I
Lengua extranjera I
Religión o Actividades de estudio</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td rowspan="4">Propias de Modalidad</td><td colspan="2">Modalidad de Artes</td><td colspan="2">Modalidad de Ciencias de la Naturaleza y de la Salud</td></tr>
<tr><td>1º</td><td>2º</td><td>1º</td><td>2º</td></tr>
<tr><td>Dibujo Artístico I
Dibujo Técnico I
Volumen</td><td>Dibujo Artístico II
Dibujo Técnico II
Fundamentos de diseño
Historia del Arte
Imagen
Técnicas de expresión gráfico-plástica</td><td>Biología y Geología
Dibujo Técnico I
Física y Química
Matemáticas I</td><td>Biología
Ciencias de la Tierra y Medioambientales
Dibujo Técnico II
Física
Matemáticas II
Química</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td rowspan="2">Optativas</td><td colspan="2">Talleres artísticos
Matemáticas de la forma
Volumen II
Ampliación de sistemas de representación técnico-gráficos</td><td colspan="2">Geología
Fisiología y Anatomía humana
Técnicas de laboratorio</td></tr>
<tr><td colspan="4">Segunda Lengua Extranjera / Música / Ciencia, Tecnología y Sociedad / Comunicación Audiovisual / Tecnología de la Información</td></tr>
</table>

<table>
<tr><td rowspan="2">Comunes</td><td>2º</td><td colspan="3">Filosofía II
Lengua Extranjera II
Historia
Lengua castellana, lengua oficial propia de las CC.AA. y Literatura II</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td rowspan="4">Propias de Modalidad</td><td colspan="2">Modalidad de Humanidades y Ciencias Sociales</td><td colspan="2">Modalidad de Tecnología</td></tr>
<tr><td>1º</td><td>2º</td><td>1º</td><td>2º</td></tr>
<tr><td>Economía
Griego I
Historia del Mundo Contemporáneo
Latín I
Matemáticas aplicadas a las CC.SS I</td><td>Economía y Org. de Empresas
Geografía
Griego II
Historia del Arte
Historia de la Música
Latín II
Matemáticas aplicadas a las CC.SS II</td><td>Dibujo Técnico I
Física y Química
Matemáticas I
Tecnología industrial I</td><td>Dibujo Técnico II
Electrotecnia
Física
Matemáticas II
Mecánica
Tecnología industrial II</td></tr>
<tr><td rowspan="2">Optativas</td><td colspan="2">Psicología
Literatura Española y Universal
Fundamentos de Administración y Gestión</td><td colspan="2">Principios fundamentales de Electrónica
Técnicas de laboratorio</td></tr>
<tr><td colspan="4">Segunda Lengua Extranjera / Música / Ciencia, Tecnología y Sociedad / Comunicación Audiovisual / Tecnología de la Información</td></tr>
</table>

BIBLIOGRAFÍA

BIBLIOGRAFÍA

1. P. Resnick and H.R. Varian, *Recommender systems*, ACM Press, 1997, p. 56-58.

2. J.B. Schafer, et al., *E-Commerce Recommender Applications*, Kluwer Academic Publishers, 2001, p. 115-153.

3. R. Burke, "Knowledge-based Recommender Systems," *Book Knowledge-based Recommender Systems*, Series Knowledge-based Recommender Systems Vol. 69, Supplement 32, ed., Editor ed.^eds., Marcel Dekker, 2000, pp.

4. D. Goldberg, et al., *Using collaborative filtering to weave an information tapestry*, ACM Press, 1992, p. 61-70.

5. G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, 2005, pp. 734-749.

6. L. Martínez, et al., "A multigranular linguistic content-based recommendation model: Research Articles," *International Journal of Intelligent Systems*, vol. 22, no. 5, 2007, pp. 419-434; DOI http://dx.doi.org/10.1002/int.v22:5.

7. D. Billsus and M.J. Pazzani, "Learning Collaborative Information Filters," *Proc. Proceedings of the 15th International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., San Francisco, CA, 1998, pp. 46-54.

8. J.L. Herlocker, et al., "An algorithmic framework for performing collaborative filtering," *Sigir'99: Proceedings of 22nd International Conference on Research and Development in Information Retrieval*, M. Hearst, et al., eds., Assoc Computing Machinery, 1999, pp. 230-237.

9. J.L. Herlocker, et al., "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, 2004, pp. 5-53.

10. D. Kim and B.J. Yum, "Collaborative filtering based on iterative principal component analysis," *Expert Syst. Appl.*, vol. 28, no. 4, 2005, pp. 823-830.

11. A. Kohrs and B. Merialdo, "Clustering for collaborative filtering applications," *Computational Intelligence for Modelling, Control & Automation - Intelligent Image Processing, Data Analysis & Information Retrieval*, Concurrent Systems Engineering Series 56, M. Mohammadian, ed., I O S Press, 1999, pp. 199-204.

12. M.J. Pazzani, "A framework for collaborative, content-based and demographic filtering," *Artif. Intell. Rev.*, vol. 13, no. 5-6, 1999, pp. 393-408.

13. L.H. Ungar and D.P. Foster, "Clustering Methods for Collaborative Filtering," *Proc. Proceedings of the Workshop on Recommendation Systems*, AAAI Press, Menlo Park California, 1998.

14. D. Billsus and M.J. Pazzani, *A hybrid user model for news story classification*, Springer-Verlag New York, Inc., 1999, p. 99-108.

15. R.J. Mooney and L. Roy, *Content-based book recommending using learning for text categorization*, ACM Press, 2000, p. 195-204.

16. M.J. Pazzani and D. Billsus, *Learning and Revising User Profiles: The Identification of Interesting Web Sites*, Kluwer Academic Publishers, 1997, p. 313-331.

17. J.D. Ullman, *Principles of database and knowledge-base systems, Vol. I*, Computer Science Press, Inc., 1988, p. 631.

18. R. Burke, et al., "Knowledge-based Navigation of Complex Information Spaces," *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, AAAI Press/MIT Press, 1996.

19. R. Burke, "Hybrid recommender systems: Survey and experiments," *User Model. User-Adapt. Interact.*, vol. 12, no. 4, 2002, pp. 331-370.

20. B. Bezerra and F.D. de Carvalho, "A symbolic hybrid approach to face the new user problem in recommender systems," *Ai 2004: Advances in Artificial Intelligence*, *Proceedings*, Lecture Notes in Artificial Intelligence 3339, Springer-Verlag Berlin, 2004, pp. 1011-1016.

21. K.Y. Jung, et al., "Hybrid collaborative filtering and content-based filtering for improved recommender system," *Computational Science - Iccs 2004, Pt 1, Proceedings*, Lecture Notes in Computer Science 3036, Springer-Verlag Berlin, 2004, pp. 295-302.

22. Y. Li, et al., "A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce," *Expert Syst. Appl.*2005; http://dblp.uni-trier.de.

23. C. Basu, et al., *Recommendation as classification: Using social and content-based information in recommendation*, American Association for Artificial Intelligence, 1998, p. 714-720.

24. T. Tran and R. Cohen, "Hybrid Recommender Systems for Electronic Commerce," *Book Hybrid Recommender Systems for Electronic Commerce*, Series Hybrid Recommender Systems for Electronic Commerce, ed., Editor ed.^eds., AAAI Press, 2000, pp.

25. Y.H. Cho and J.K. Kim, "Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce," *Expert Syst. Appl.*, vol. 26, 2004, pp. 223-246; DOI 10.1016/S0957-4174(03)00138-6.

26. J.L. Herlocker, et al., "Explaining collaborative filtering recommendations," *Proc. Proceedings of the 2000 ACM conference on Computer supported cooperative work*, ACM Press, 2000, pp. 241-250.

27. K. Yu, et al., "Instance Selection Techniques for Memory-Based Collaborative Filtering," *Proc. Proceedings of 2nd SIAM International Conference on Data Mining*, SIAM, 2002.

28. M. Eirinaki and M. Vazirgiannis, "Web Mining for Web Personalization," *ACM Transactions on Internet Technology*, vol. 3, no. 1, 2003, pp. 1-27.

29. M.M. Recker, et al., "What do you recommend? Implementation and analyses of collaborative information filtering of Web resources for education," *Instr. Sci.*, vol. 31, no. 4-5, 2003, pp. 299-316.

30. C. O'Riordan and J. Griffith, "Providing personalised recommendations in a web-based education system," *Knowledge-Based Intellignet Information and Engineering Systems, Pt 2, Proceedings*, Lecture Notes in Artificial Intelligence 2774, Springer-Verlag Berlin, 2003, pp. 245-251.

31. J.A. Oravec, "Collaborative filtering applications in distance education," *17th Annual Conference on Distance Teaching and Learning, Conference Proceedings*, University Wisconsin Distance Teaching & Learning Conference, 2001, pp. 281-284.

32. G. Liang, et al., "Courseware Recommendation in E-Learning System," *Lecture Notes in Computer Science*, vol. 4181, 2006, pp. 10-24.

33. M. Khine and A. Lourdusamy, "Blended learning approach in teacher education: combining face-to-face instruction, multimedia viewing and online discussion," *British Journal of Educational Technology*, vol. 34, no. 5, 2003, pp. 671-675.

34. L. Reis, et al., "Distance learning as a tool to support a classroom based learning: College of business administration challenge," *INFORMATION TECHNOLOGY AND ORGANIZATIONS: TRENDS, ISSUES, CHALLENGES AND SOLUTIONS, VOLS 1 AND 2*, M. KhosrowPour, ed., 2003, pp. 1126, 1128.

35. , *1º Cuaderno de Orientación. Educación Secundaria Obligatoria.*, Equipo de Orientación Educativa, Junta de Andalucía.

36. , *2º Cuaderno de Orientación. Educación Secundaria Obligatoria.*, Equipo de Orientación Educativa, Junta de Andalucía.

37. *Decreto 148/2002 de 14 de mayo de 2002*.

38. *Decreto por el que se establecen las enseñanzas correspondientes al Bachillerato en Andalucía*.

39. *Orden por la que se establece el horario lectivo, las materias propias de la modalidad, las materias optativas y los itinerarios educativos correspondientes al Bachillerato.*

40. , *Orientación profesional: Programa Elige.*, Instituto Andaluz de la Mujer. Junta de Andalucía.

41. J.S. Breese, et al., "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," *Proc. Uncertainty in Artificial Intelligence. Proceedings of the Fourteenth Conference.*, 1998, pp. 18.

42. M.P. O'Mahony, et al., "An evaluation of neighbourhood formation on the performance of collaborative filtering," *Artif. Intell. Rev.*, vol. 21, no. 3-4, 2004, pp. 215-228.

43. P. Resnick, et al., "GroupLens: an open architecture for collaborative filtering of netnews," *Proc. Proceedings of the 1994 ACM conference on Computer supported cooperative work*, ACM Press, 1994, pp. 175-186.

44. J.A. Konstan, et al., "GroupLens: Applying collaborative filtering to Usenet news," *Communications of the Acm.*, vol. Vol. 40, no. 3, 1997, pp. 77-87.

45. W. Hill, et al., "Recommending and Evaluating Choices in a Virtual Community of Use.," *Proc. Proceedings of ACM CHI'95 Conference on human factors in computing systems.*, 1995, pp. 194-201.

46. U. Shardanand and P. Maes, "Social Information Filtering: Algorithms for Automating "Word of Mouth"." *Proc. Proceedings of ACM CHI '95*, 1995, pp. 210-217.

47. U. Wolz, et al., *Computer-mediated communication in collaborative educational settings (report of the ITiCSE '97 working group on CMC in collaborative educational settings)*, ACM Press, 1997, p. 51-69.

48. B. Sarwar, et al., *Analysis of recommendation algorithms for e-commerce*, ACM Press, 2000, p. 158-167.

49. Z.Q. Wang and B.Q. Feng, "Collaborative filtering algorithm based on mutual information," *Advanced Web Technologies and Applications*, Lecture Notes in Computer Science 3007, Springer-Verlag Berlin, 2004, pp. 405-415.

50. M. Deshpande and G. Karypis, "Item-based top-N recommendation algorithms," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, 2004, pp. 143-177.

51. C. Zeng, et al., "Similarity measure and instance selection for collaborative filtering," *Int. J. Electron. Commer.*, vol. 8, no. 4, 2004, pp. 115-129.

52. R. Jin, et al., *An automatic weighting scheme for collaborative filtering*, ACM Press, 2004, p. 337-344.

53. G.-R. Xue, et al., *Scalable collaborative filtering using cluster-based smoothing*, ACM Press, 2005, p. 114-121.

54. M.K. Condliff, et al., "Bayesian Mixed-Effects Models for Recommender Systems," *Proc. Proceedings of the ACM SIGIR Workshop on Recommender Systems: Algorithms and Evaluation. 22nd Intl. Conf. on Research and Development in Information Retrieval*, 1999.

55. S.M. Weiss and C.A. Kulikowski, *Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems*, Morgan Kaufmann Publishers Inc., 1991, p. 223.

56. N.J. Belkin and W.B. Croft, *Information filtering and information retrieval: two sides of the same coin?*, ACM Press, 1992, p. 29-38.

57. M.J. Pazzani and D. Billsus, *Learning and Revising User Profiles: The Identification ofInteresting Web Sites*, Kluwer Academic Publishers, 1997, p. 313-331.

58. B. Sarwar, et al., *Item-based collaborative filtering recommendation algorithms*, ACM Press, 2001, p. 285-295.

59. K. Yu, et al., *Selecting relevant instances for efficient and accurate collaborative filtering*, ACM Press, 2001, p. 239-246.

60. Y. Kai, et al., "Instance selection techniques for memory-based collaborative filtering," *Proceedings of the Second Siam International Conference on Data Mining*, Siam Proceedings Series, R. Grossman, et al., eds., Siam, 2002, pp. 59-74.

61. T.-H. Kim and S.-B. Yang, *An Effective Recommendation Algorithm for Improving Prediction Quality*, Springer Berlin / Heidelberg, 2006, p. 1288-1292.

62. T.H. Kim and S.B. Yang, "Using attributes to improve prediction quality in collaborative filtering," *E-Commerce and Web Technologies*, Lecture Notes in Computer Science 3182, Springer-Verlag Berlin, 2004, pp. 1-10.

63. M.P. O'Mahony, et al., "Towards robust collaborative filtering," *Artificial Intelligence and Cognitive Science, Proceedings*, Lecture Notes in Artificial Intelligence 2464, Springer-Verlag Berlin, 2002, pp. 87-94.

64. J. Lim, "ADOdb Library for PHP," 2006; http://phplens.com/lens/adodb/docs-adodb.htm.

65. M. Achour, et al., "PHP Manual," *Book PHP Manual*, Series PHP Manual, ed., Editor ed.^eds., 2007, pp.

66. "MySQL 5.1 Reference Manual," 2007; http://dev.mysql.com/doc/refman/5.1/en/index.html.

67. D. Raggett, et al., "HTML 4.01 Specification," 1999; http://www.w3.org/TR/html4/.

68. B. Bos, et al., “Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification,” 2007; http://www.w3.org/TR/2007/CR-CSS21-20070719/.

69. H.W. Lie and B. Bos, “Cascading Style Sheets, level 1,” 1999; http://www.w3.org/TR/1999/REC-CSS1-19990111.

70. J. Canny, “Collaborative filtering with privacy,” *2002 Ieee Symposium on Security and Privacy, Proceedings*, Proceedings: Ieee Symposium on Security and Privacy, Ieee Computer Soc, 2002, pp. 45-57.

71. W.F.J. Verhaegh, et al., “Privacy protection in memory-based collaborative filtering,” *Ambient Intelligence, Proceedings*, Lecture Notes in Computer Science 3295, Springer-Verlag Berlin, 2004, pp. 61-71.

72. H. Polat and W.L. Du, “Privacy-preserving collaborative filtering,” *Int. J. Electron. Commer.*, vol. 9, no. 4, 2005, pp. 9-35.