

Appendix B

Introducción

En este capítulo presentamos una visión general de la tesis. En las Secciones B.1 y B.2 mostramos las motivaciones y objetivos de nuestro trabajo. En la Sección B.3 resumimos las principales contribuciones de la tesis, y en la Sección B.4 enumeramos las publicaciones resultantes de nuestra investigación. Finalmente, en la Sección B.5 describimos la estructura de este documento.

B.1 Motivación

Las tecnologías de Recuperación de Información (RI) han ganado una prevalencia excepcional en las dos últimas décadas con la explosión del número de repositorios masivos de información existentes en línea, y más en particular en la *World Wide Web*. En RI se han investigado y diseñado formas que buscan maximizar el grado de satisfacción de ciertas condiciones objetivas, típicamente – aunque no necesariamente de manera única – la satisfacción del usuario. La investigación y el desarrollo en RI han girado en torno a la definición de modelos y algoritmos que mejor alcancen dicho objetivo, de metodologías y métricas que permiten evaluar cuánto de bien se consigue esta meta con diferentes sistemas, así como de teorías consolidadas que proveen una base sólida y una orientación en el desarrollo de algoritmos de RI y su consistente evaluación. Entre las muchas tendencias que han surgido desde el flujo principal de investigación y desarrollo, un nuevo reto de investigación ha empezado a ser considerado desde comienzos de los 2000: ¿es posible predecir cómo de bueno será un resultado devuelto por un sistema de RI antes de presentarlo al usuario, o incluso, antes de efectuar por completo la búsqueda del resultado por el sistema (Cronen-Townsend et al., 2002)? Esta pregunta ha dado pie a una fértil corriente de investigación en lo que se ha llamado en RI como **predicción de eficacia**.

La predicción de eficacia tiene muchos usos potenciales en RI. Desde la perspectiva del usuario proporcionaría información que puede ser usada para dirigir una búsqueda, desde la perspectiva del sistema ayudaría a distinguir consultas poco eficaces, y desde la perspectiva del administrador del sistema permitiría identificar consultas relacionadas sobre un tema específico que resultan difíciles para el motor de búsqueda. Las técnicas de predicción de eficacia se basan en el análisis y caracterización de la evidencia usada por un sistema de RI para evaluar la relevancia (utilidad, valor, etc.) de los ítems a recuperar (documentos, artículos, etc.) en tiempo de ejecución (Cronen-Townsend et al., 2002). El escenario más común en recuperación de información supone una consulta del usuario y una colección de documentos como la entrada básica para formar una lista ordenada de resultados de búsqueda, pero otros elementos adicionales pueden tenerse en cuenta para seleccionar y ordenar resultados (Baeza-Yates and Ribeiro-Neto, 2011). Cualquier información que el sistema de recuperación de información tome como entrada puede ser utilizada también para la predicción de su eficacia, y habitualmente los métodos de predicción usan información adicional. El contexto del usuario (la tarea actual, los registros de búsquedas, las preferencias, etc.), las propiedades globales de los documentos de la colección, comparaciones con respecto a (otros) elementos de referencia como datos históricos, o la salida de distintos sistemas, son algunos ejemplos de las diferentes fuentes de información de las que un predictor puede extraer evidencia.

Predecir la eficacia de un subsistema, módulo, función o entrada contrastando la estimación de eficacia de cada componente para una consulta, permite una serie de estrategias de optimización dinámicas que seleccionen en tiempo de ejecución la opción que se predice funcionará mejor o, cuando se usan sistemas a gran escala o aproximaciones híbridas, que ajusten sobre la marcha el grado de participación de cada módulo. En el campo de RI dominan los casos donde información de relevancia, sistemas de recuperación, modelos y criterios se definen en función de la fusión o combinación de sub-modelos. Sistemas personalizados de recuperación (incluyendo técnicas como búsqueda personalizada, sistemas de recomendación, filtrado colaborativo y búsqueda contextualizada) son claros ejemplos donde se puede aplicar la predicción de eficacia dado que dichos sistemas combinan varias fuentes de evidencia para la estimación de la relevancia, como pueden ser consultas explícitas, historial de búsqueda, puntuaciones de usuarios, información social, información del usuario y modelos de contexto.

La predicción de eficacia encuentra una motivación adicional en la recomendación personalizada, puesto que esas aplicaciones pueden decidir si producir recomendaciones u ocultarlas, entregando sólo las suficientemente fiables. Más aún, los Sistemas de Recomendación (SR) actuales se caracterizan por una creciente diversificación de los tipos y fuentes de datos, contenidos, evidencias y métodos disponibles para tomar decisiones y construir los resultados. En este contexto, predecir la eficacia de un método de recomendación específico o de una componente se convierte en un problema atractivo, ya que permite una combinación adecuada de las alternativas disponibles y sacar el máximo provecho de ellas, adaptando dinámicamente la estrategia de recomendación a la situación actual. La cuestión gana mayor relevancia hoy con la proliferación de técnicas de recomendación híbridas para mejorar la precisión de los métodos – siendo el premio Netflix uno de los ejemplos paradigmáticos del uso de estos métodos, donde los participantes mejor situados combinaron múltiples métodos de recomendación. Esto requiere de una investigación en aproximaciones híbridas con un nivel de mecanismos dinámicos auto-ajustables, de manera que se optimice la efectividad resultante de los sistemas de recomendación, tomando oportuna ventaja de datos de alta calidad cuando estén disponibles, pero evitando aferrarse a estrategias fijas cuando se predice que pueden producir resultados pobres bajo ciertas condiciones.

La predicción de eficacia en RI típicamente se evalúa en términos de correlación entre el predictor y los valores de eficacia para cada consulta. Esto requiere métricas de evaluación de la eficacia fiables, las cuales han sido analizadas cuidadosamente y están actualmente bien establecidas en el área de RI, orientadas en su mayor parte a búsqueda ad-hoc. Por el contrario, la evaluación en el campo de los SR está más abierta, y la variabilidad en las técnicas de evaluación y configuraciones experimentales es significativa. Cómo medir la eficacia de un sistema de recomendación es un

asunto clave en nuestra investigación ya que las medidas de calidad del sistema pueden verse influidas por propiedades estadísticas del método de medición y/o por el diseño experimental. A lo largo de esta tesis nos centraremos en la precisión del sistema, donde hemos de evitar que si una métrica estuviera sesgada hacia algún tipo de ruido medido a la vez que la calidad de la recomendación, pues entonces un predictor que sólo capturara dicho ruido podría actuar como un predictor de eficacia erróneamente útil. Así, los sesgos estadísticos (ruidos) de las metodologías de evaluación deben ser bien entendidos para permitir una valoración significativa de los predictores de eficacia.

Tomando el estado del arte de predicción de eficacia en RI como punto de partida, el trabajo actual replantea este problema en el campo de los Sistemas de Recomendación donde ha sido escasamente considerado hasta la fecha. Investigamos definiciones adecuadas de eficacia en el contexto de los SR y los elementos a los que sensatamente se puede aplicar, analizando los sesgos estadísticos que pueden aparecer cuando se adapta el marco de evaluación de RI a SR. De este modo tomamos como dirección principal la aplicación de los predictores de eficacia para obtener mejoras en dos problemas específicos de combinación en el campo de SR, a saber, la combinación dinámica de métodos de recomendación en sistemas de recomendación híbridos, y la agregación dinámica de las señales de vecinos en filtrado colaborativo basado en usuario.

B.2 Objetivos

El principal objetivo de la investigación presentada aquí es encontrar métodos predictivos para la eficacia de componentes específicos de sistemas de recomendación, y mejorar la eficacia de los métodos de recomendación combinados, basados en el análisis y predicción dinámicos y automáticos de la eficacia esperada de los elementos de un método compuesto, con los cuales la participación relativa de cada elemento se ajusta de acuerdo a su efectividad predicha. Para abordar estos problemas nuestro trabajo tiene los siguientes objetivos de investigación concretos:

O1: Análisis y formalización de cómo se define y evalúa la eficacia en los sistemas de recomendación. Dado que pretendemos predecir su eficacia necesitamos desarrollar un estudio en profundidad sobre cómo se pueden evaluar de manera fiable los sistemas de recomendación en términos de valores numéricos de una métrica. Más aún, hemos de investigar si existen sesgos en la manera en que los sistemas se evalúan – debidos tanto a metodologías de evaluación como a métricas, ya que cualquier sesgo en el proceso de evaluación podría conducir a resultados inconcluyentes o engañosos con respecto al poder predictivo de los métodos de predicción de eficacia propuestos. Si dichos sesgos existieran, intentaríamos entenderlos de manera precisa, y desarrollar metodologías que los aislaran. Además, deberíamos com-

probar la efectividad de nuestros predictores frente a la de otros métodos conocidos más básicos y observar si cambia al aislar dichos sesgos.

O2: Adaptación y definición de técnicas de predicción de eficacia para sistemas de recomendación. Queremos estudiar el potencial de la predicción de eficacia en problemas y escenarios específicos en el área de los Sistemas de Recomendación. Investigaremos la definición de un marco formal donde los predictores de eficacia puedan ser integrados. Como punto de partida, pretendemos explorar la adaptación de predictores específicamente efectivos en Recuperación de Información como la claridad de la consulta (Cronen-Townsend et al., 2002) al campo de la recomendación. De manera complementaria a la adaptación de técnicas conocidas, queremos investigar la definición de nuevos predictores basados en modelos de la Teoría de Información y Grafos Sociales, además de otras aproximaciones heurísticas y específicas del dominio. Una vez hayamos definido algunos predictores de eficacia para recomendación, evaluaríamos la efectividad de dichos predictores en términos de su correlación con métricas de eficacia de manera que pudiéramos estimar su poder de predicción.

O3: Aplicación de predictores de eficacia a sistemas de recomendación compuestos e híbridos. Queremos identificar e integrar los predictores propuestos en métodos de recomendación combinados para obtener una mejora real en la eficacia de los métodos combinados. Con este objetivo en mente consideraremos problemas donde la agregación de métodos de recomendación es necesaria, y analizaremos cómo aplicar los predictores de eficacia mencionados antes a tales problemas. Además, deberíamos realizar un estudio metodológico para la aproximación experimental, su configuración y las métricas usadas, de manera que se usen métodos base y diseños experimentales apropiados. Finalmente, evaluaremos las mejoras y beneficios de los métodos combinados cuando se utilizan los predictores de eficacia.

B.3 Contribuciones

Esta tesis se dedica al problema de estimar la eficacia de los sistemas de recomendación para usuarios e ítems particulares. Las contribuciones particulares de esta tesis están relacionados con la evaluación de la eficacia de un sistema de recomendación y la predicción de la misma, donde hemos abordado varios problemas relacionados con ambos temas y hemos propuesto modelos y métodos novedosos, que han sido empleados en dos aplicaciones como mostraremos a continuación.

Como un primer paso, esta tesis analiza el paradigma Cranfield de evaluación de Recuperación de Información, dado que los sistemas de recomendación normalmente se consideran como un problema particular del filtrado de información, y, por tanto, de la recuperación de información en general (Belkin and Croft, 1992). En el Capítulo 4 **argumentamos las diferencias involucradas en las alternativas de**

diseños experimentales a partir de las hipótesis habituales hechas en el paradigma Cranfield, lo cual resulta en **la aparición de sesgos estadísticos considerables en Sistemas de Recomendación**, para los que **proponemos diferentes métodos para neutralizar estos sesgos**. De manera adicional, las siguientes contribuciones relacionadas han sido realizadas:

- Proponemos una caracterización precisa y sistemática de las alternativas para la adaptación del paradigma Cranfield a tareas de recomendación. Identificamos hipótesis y condiciones subyacentes en dicho paradigma Cranfield que no pueden ser asumidas en los experimentos habituales de recomendación.
- Detectamos y caracterizamos los sesgos estadísticos resultantes, a saber, la dispersión de test y la popularidad de los ítems, los cuales no aparecen en colecciones de test habituales en RI, pero que interfieren en experimentos de recomendación.
- Proponemos dos diseños experimentales nuevos para neutralizar estos sesgos. Observamos que una evaluación basada en percentiles reduce considerablemente el margen para el sesgo de popularidad, mientras que una aproximación basada en un test uniforme elimina cualquier ventaja estadística obtenida por tener más puntuaciones positivas de test. Más aún, encontramos que las dos propuestas discriminan bien entre recomendaciones puramente basadas en popularidad y un algoritmo de recomendación personalizado.

Además, en esta tesis **mostramos cómo las técnicas de predicción de eficacia de consultas desarrolladas en Recuperación de Información pueden ser adaptadas a los Sistemas de Recomendación, y resultar en predictores eficaces en este dominio**. Presentamos estos predictores de eficacia en el Capítulo 6, donde proponemos distintas adaptaciones del predictor de claridad de consulta basadas en distintas interpretaciones de los modelos de lenguaje subyacentes, así como con modelos de Teoría de Información y de Grafos Sociales. Más aún, en el mismo capítulo **evaluamos la efectividad de dichos predictores midiendo la correlación con respecto a métricas de eficacia**, donde también probamos los métodos propuestos en el Capítulo 4 para neutralizar los sesgos en la evaluación. A continuación resumimos las contribuciones específicas respecto a la predicción de eficacia para recomendación:

- Definimos y elaboramos varios modelos predictivos en el dominio de los sistemas de recomendación de acuerdo a diferentes formulaciones e hipótesis, y basados en tres tipos de datos de preferencia: puntuaciones, registros y sociales.
- Las formulaciones para preferencias en base a puntuaciones se basan en adaptaciones de la claridad de consulta de RI y conceptos de Teoría de la Información como la entropía. En esta adaptación proponemos distintas estimaciones

de probabilidad, donde desarrollamos derivaciones Bayesianas y estimaciones no paramétricas.

- También explotamos atributos temporales al definir los predictores basados en registros. Más específicamente, usamos una versión sensible al tiempo de la divergencia de Kullback-Leibler, junto con otros conceptos de series temporales como la autocorrelación del usuario.
- Usamos métricas basadas en Teoría de Grafos para definir predictores que aprovechan las estructuras de red social y las correlaciones entre las propiedades topológicas de los usuarios y el éxito de las recomendaciones devueltas.
- Encontramos fuertes correlaciones entre las salidas de los predictores y las métricas de eficacia, mostrando por tanto evidencia empírica del poder predictivo de las técnicas propuestas. Más aún, cuando se utilizan las metodologías no sesgadas los predictores conservan buenos valores de correlación, evidenciando que los predictores propuestos no están sólo capturando los sesgos analizados y beneficiándose de los mismos, especialmente cuando los comparamos frente a otros predictores triviales.

Finalmente, los Capítulos 7 y 8 presentan dos aplicaciones de los predictores de eficacia en Sistemas de Recomendación. En el Capítulo 7 **proponemos varios sistemas híbridos ponderados linealmente donde las ponderaciones se ajustan dinámicamente de acuerdo a la salida de los predictores**. Observamos que las correlaciones obtenidas en el Capítulo 6 ayudan a decidir cuáles son las mejores combinaciones a experimentar. Más importante aún, **la correlación entre el predictor y el algoritmo de recomendación tiende a anticipar bien cuándo un sistema híbrido mejorará con respecto a un método base**. Además, el Capítulo 8 **presenta un marco unificado donde los predictores de eficacia se usan para seleccionar y ponderar los vecinos cercanos en un algoritmo estándar de filtrado colaborativo basado en usuario**. La metodología tradicional de predicción de eficacia es adaptada y traducida a este problema, donde definimos novedosas métricas sobre la eficacia de vecinos y evaluamos el poder predictivo de los predictores.

Las contribuciones relacionadas con la parte de aplicaciones de la tesis son, en resumen, las siguientes:

- Proponemos un marco de hibridación dinámica para decidir automáticamente cuándo y cómo debería realizarse la hibridación, dependiendo de distintas condiciones, a saber: las correlaciones entre los algoritmos de recomendación y los predictores, y la eficacia relativa entre los algoritmos combinados.
- En varios experimentos con los predictores de eficacia mencionados previamente, nuestros resultados indican que una fuerte correlación con la eficacia

tiende a corresponder con mejoras en la recomendación híbrida dinámica cuando los predictores se usan para ajustar los pesos de la combinación.

- Proponemos un marco teórico para la selección y ponderación de vecinos en sistemas de recomendación basados en usuarios. Este marco se fundamenta en la predicción de eficacia estableciendo equivalencias entre la tarea de predicción de puntuaciones basada en vecindarios y una agregación dinámica de componentes, en este caso, las predicciones de cada vecino.
- Comparamos varias métricas de confianza del estado del arte así como otras técnicas para valorar vecinos, interpretadas ambas como predictores de eficacia de vecinos. También proponemos varias métricas de eficacia de vecinos que capturan diferentes nociones de la calidad de los vecinos.

B.4 Publicaciones relacionadas con la tesis

En los siguientes artículos de revistas y conferencias internacionales presentamos descripciones, resultados y conclusiones relacionadas con esta tesis:

Predicción de eficacia y evaluación

1. Bellogín, A., Cantador, I., Díez, F., Castells, P., and Chavarriaga, E. (2012). An empirical comparison of social, collaborative filtering, and hybrid recommenders. *ACM Transactions on Intelligent Systems and Technology*, por aparecer.
2. Bellogín, A., Castells, P., and Cantador, I. (2011). Predicting the Performance of Recommender Systems: An Information Theoretic Approach. In Amati, G. and Crestani, F., editors, *ICTIR*, volume 6931 of *Lecture Notes in Computer Science*, pages 27–39, Berlin, Heidelberg. Springer Berlin / Heidelberg.
3. Bellogín, A., Castells, P., and Cantador, I. (2011). Self-adjusting hybrid recommenders based on social network analysis. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information*, SIGIR '11, pages 1147–1148, New York, NY, USA. ACM.
4. Bellogín, A., Castells, P., and Cantador, I. (2011). Precision-oriented evaluation of recommender systems: an algorithmic comparison. In *Proceedings of the fifth ACM conference on Recommender systems*, RecSys '11, pages 333–336, New York, NY, USA. ACM.
5. Bellogín, A. and Castells, P. (2010). A Performance Prediction Approach to Enhance Collaborative Filtering Performance. In Gurrin, C., He, Y., Kazai, G., Kruschwitz, U., Little, S., Roelleke, T., Rüger, S., and Rijsbergen, editors,

Advances in Information Retrieval, volume 5993 of *Lecture Notes in Computer Science*, pages 382–393, Berlin, Heidelberg. Springer Berlin / Heidelberg.

6. Bellogín, A. and Castells, P. (2009). Predicting neighbor goodness in collaborative filtering. In And, T. A., Yager and, R. R., And, H. B., And, H. C., and Larsen, H. L., editors, *FQAS*, volume 5822 of *Lecture Notes in Computer Science*, pages 605–616, Berlin, Heidelberg. Springer Berlin / Heidelberg.

Recomendación basada en contenido

7. Cantador, I., Bellogín, A., and Vallet, D. (2010). Content-based recommendation in social tagging systems. In *Proceedings of the fourth ACM conference on Recommender systems*, RecSys '10, pages 237–240, New York, NY, USA. ACM.
8. Cantador, I., Bellogín, A., and Castells, P. (2008). News@hand: A Semantic Web Approach to Recommending News. In Nejd, W., Kay, J., Pu, P., and Herder, E., editors, *Adaptive Hypermedia and Adaptive Web-Based Systems*, volume 5149 of *Lecture Notes in Computer Science*, chapter 34, pages 279–283. Springer Berlin / Heidelberg, Berlin, Heidelberg.
9. Cantador, I., Bellogín, A., and Castells, P. (2009). Ontology-Based Personalised and Context-Aware Recommendations of News Items. In *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on*, volume 1, pages 562–565.
10. Cantador, I., Bellogín, A., Fernández-Tobías, I., and López-Hernández, S. (2011a). Semantic Contextualisation of Social Tag-Based Profiles and Item Recommendations. In Huemer, C., Setzer, T., Aalst, W., Mylopoulos, J., Sadeh, N. M., Shaw, M. J., Szyperski, C., Aalst, W., Mylopoulos, J., Sadeh, N. M., Shaw, M. J., and Szyperski, C., editors, *Electronic Commerce and Web Technologies*, volume 85 of *Lecture Notes in Business Information Processing*, chapter 9, pages 101–113. Springer Berlin Heidelberg, Berlin, Heidelberg.
11. Fernández-Tobías, I., Cantador, I., and Bellogín, A. (2011). cTag: Semantic contextualisation of social tags. In *Proceedings of the Workshop on Semantic Adaptive Social Web (SASWeb 2011)*. *CEUR Workshop Proceedings, vol. 730*, pages 45–54. RWTH, Aachen (2011).

Recomendación basada en filtrado colaborativo

12. Bellogín, A., Wang, J., and Castells, P. Bridging Memory-Based Collaborative Filtering and Text Retrieval. *Information Retrieval Journal*, por aparecer.
13. Bellogín, A., Cantador, I., and Castells, P. A Comparative Study of Heterogeneous Item Recommendations in Social Systems. *Information Sciences*, por aparecer.

14. Bellogín, A. and Parapar, J. (2012). Using Graph Partitioning Techniques for Neighbour Selection in User-Based Collaborative Filtering. In *Proceedings of the sixth ACM conference on Recommender systems*, RecSys '12, pages 213–216, New York, NY, USA. ACM. (premio al mejor artículo corto)
15. Bellogín, A., Wang, J., and Castells, P. (2011). Structured collaborative filtering. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 2257–2260, New York, NY, USA. ACM.
16. Bellogín, A., Wang, J., and Castells, P. (2011). Text Retrieval Methods for Item Ranking in Collaborative Filtering. In Clough, P., Foley, C., Gurrin, C., Jones, G., Kraaij, W., Lee, H., and Mudoch, V., editors, *Advances in Information Retrieval*, volume 6611 of *Lecture Notes in Computer Science*, chapter 30, pages 301–306. Springer Berlin / Heidelberg, Berlin, Heidelberg.
17. Bellogín, A., Cantador, I., and Castells, P. (2010). A study of heterogeneity in recommendations for a social music service. In *Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems*, HetRec '10, pages 1–8, New York, NY, USA. ACM.

Recomendación basada en filtrado social

18. Díez, F., Chavarriaga, J. E., Campos, P. G., and Bellogín, A. (2010). Movie recommendations based in explicit and implicit features extracted from the Filmtipset dataset. In *Proceedings of the Workshop on Context-Aware Movie Recommendation*, CAMRa '10, pages 45–52, New York, NY, USA. ACM.

Recomendación sensible al tiempo

19. Campos, P. G., Bellogín, A., Díez, F., and Cantador, I. (2012). Time Feature Selection for Identifying Active Household Members. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, CIKM '12, New York, NY, USA. ACM (por aparecer).
20. Campos, P. G., Díez, F., and Bellogín, A. (2011). Temporal rating habits: a valuable tool for rating discrimination. In *Proceedings of the 2nd Challenge on Context-Aware Movie Recommendation*, CAMRa '11, pages 29–35, New York, NY, USA. ACM.
21. Campos, P. G., Bellogín, A., Díez, F., and Chavarriaga, J. E. (2010). Simple time-biased KNN-based recommendations. In *Proceedings of the Workshop on Context-Aware Movie Recommendation*, CAMRa '10, pages 20–23, New York, NY, USA. ACM.

Sistemas de recomendación híbridos

22. Cantador, I., Castells, P., and Bellogín, A. (2011). An enhanced semantic layer for hybrid recommender systems. *International Journal on Semantic Web and Information Systems*, 7(1):44–78.
23. Cantador, I., Bellogín, A., and Castells, P. (2008). A multilayer ontology-based hybrid recommendation model. *AI Commun.*, 21(2-3):203–210.
24. Cantador, I., Castells, P., and Bellogín, A. (2007). Modelling Ontology-based Multilayered Communities of Interest for Hybrid Recommendations. In *Workshop on Adaptation and Personalisation in Social Systems: Groups, Teams, Communities, at the 11th International Conference on User Modeling*.

Evaluación de la recomendación

25. Bellogín, A., Cantador, I., Castells, P., and Ortigosa, A. (2011). Discerning Relevant Model Features in a Content-based Collaborative Recommender System. In Fürnkranz, J. and Hüllermeier, E., editors, *Preference Learning*, chapter 20, pages 429–455. Springer Berlin Heidelberg, Berlin, Heidelberg.
26. Bellogín, A., Cantador, I., Castells, P., and Ortigosa, A. (2008). Discovering Relevant Preferences in a Personalised Recommender System using Machine Learning Techniques. In *Preference Learning Workshop (PL 2008), at the 8th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2008)*, pages 82–96.

Estas publicaciones se relacionan con los contenidos de esta tesis como sigue. En [4] analizamos diferentes metodologías de evaluación existentes en la literatura de recomendación (Capítulos 3 y 4). En [2], [5], [6] definimos las formulaciones para el concepto de claridad de usuario basada en puntuaciones (Capítulo 6), mientras que en [1] y [3] definimos los predictores sociales (también en el Capítulo 6). Además, en [1] y [3] también investigamos el uso de los predictores de eficacia para recomendaciones híbridas dinámicas (Capítulo 7). Más aún, en [5] y [6] abordamos el problema de la ponderación de vecinos basada en los predictores de eficacia de vecinos (Capítulo 8).

Además, durante el transcurso de la tesis, la investigación presentada aquí motivó una serie de publicaciones que abordaban temas más amplios en el área, como las recomendaciones basadas en contenido [7-11], el filtrado colaborativo [12-17], las técnicas de filtrado social [18], la recomendación sensible al tiempo [19-21], los sistemas de recomendación híbrida [22-24], y la evaluación de la recomendación [25, 26]. Estas publicaciones han resultado en el uso y construcción de conjuntos de datos, el desarrollo de algoritmos y la investigación y utilización de algunas metodologías y métricas de evaluación que aparecen en esta tesis.

Publicaciones adicionales

Trabajo preliminar relacionado con las propuestas presentadas aquí fue incluido en primer lugar en la tesis de Máster titulada “Performance prediction in recommender Systems: Application to the dynamic optimisation of aggregative methods” (Bellogín, 2009). Específicamente, en dicho trabajo proponemos el concepto de predicción de eficacia para recomendación. Además, la motivación, el potencial impacto y los principales resultados de nuestra investigación fueron publicados como contribuciones en dos simposios doctorales internacionales:

- Bellogín, A. (2011). Predicting performance in recommender systems. Doctoral Symposium. In *Proceedings of the fifth ACM conference on Recommender systems*, Rec-Sys '11, pages 371–374, New York, NY, USA. ACM.
- Bellogín, A. (2011). Performance Prediction in Recommender Systems. Doctoral Symposium. In Konstan, J., Conejo, R., Marzo, J., and Oliver, N., editors, *User Modeling, Adaption and Personalization*, volume 6787 of *Lecture Notes in Computer Science*, pages 401–404, Berlin, Heidelberg. Springer Berlin / Heidelberg.

Además, los siguientes artículos están bajo revisión, algunos de ellos altamente relacionados con los temas de la tesis:

- Bellogín, A., Castells, P., and Cantador, I. Statistical Biases in IR Metrics for Recommender Systems: A Methodological Framework for the Adaptation of the Cranfield Paradigm. En revisión.
- Bellogín, A., Castells, P., and Cantador, I. Neighbour Selection and Weighting in User-Based Recommender Systems: A Performance Prediction Approach. En revisión.
- Parapar, J., Bellogín, A., Castells, P., and Barreiro, Á. Relevance-Based Language Modelling for Recommender Systems. En revisión.

B.5 Estructura de la tesis

La tesis está dividida en seis partes. La primera parte presenta y motiva el problema abordado, así como una revisión del estado del arte en el campo de los Sistemas de Recomendación donde esta tesis se enmarca. La segunda parte describe las diferentes técnicas de evaluación usadas en la literatura de sistemas de recomendación, y provee un análisis de las alternativas de diseño y los sesgos estadísticos que pueden aparecer. La tercera parte hace una revisión de la literatura en predicción de eficacia, propone adaptaciones de este concepto al espacio de los sistemas de recomendación, y evalúa el poder predictivo de estas propuestas. La cuarta parte muestra dos aplicaciones de los predictores de eficacia propuestos. La quinta parte concluye y resume las princi-

pales contribuciones de esta tesis. Información adicional y otros detalles se incluyen en la última parte.

Más detalladamente, los contenidos de esta tesis se distribuyen como sigue:

Parte I. Introducción

- El **Capítulo 1** presenta la motivación, objetivos, contribuciones y publicaciones relacionadas con la tesis.
- El **Capítulo 2** presenta una visión general del estado del arte en sistemas de recomendación, teniendo en cuenta una clasificación de los principales tipos de técnicas. También describe los puntos débiles de las distintas técnicas de recomendación y presenta una amplia gama de métodos híbridos de recomendación que ayudan a superar esas limitaciones.

Parte II. Evaluando la Eficacia en los Sistemas de Recomendación

- El **Capítulo 3** describe las principales métricas y metodologías de evaluación usadas en el campo de los sistemas de recomendación. También describe los conjuntos de datos públicos más habitualmente usados.
- El **Capítulo 4** presenta un análisis y formalización de las diferentes metodologías de evaluación descritas en la literatura. Primero, presenta una caracterización sistemática de las alternativas de diseños experimentales. Después identifica y analiza sesgos estadísticos que aparecen cuando algunas metodologías se aplican a recomendación, y propone dos diseños experimentales alternativos que neutralizan tales sesgos satisfactoriamente.

Parte III. Prediciendo la Eficacia en los Sistemas de Recomendación

- El **Capítulo 5** presenta el problema de predicción de eficacia en Recuperación de Información, incluye un resumen de los principales trabajos en dicha área, tanto en la definición de predictores de eficacia (aplicados a consultas) como en la evaluación de los predictores para estimar su poder predictivo.
- El **Capítulo 6** formula el problema de predicción de eficacia en los Sistemas de Recomendación. Define varios predictores de eficacia basados en tres dimensiones de la recomendación donde analizamos de manera cualitativa el poder predictivo de los predictores.

Parte IV. Aplicaciones

- El **Capítulo 7** propone un marco donde los predictores de eficacia se usan para construir sistemas de recomendación híbridos dinámicos. Evalúa estos algoritmos de recomendación en los tres espacios de entrada previamente considerados en la definición de los predictores de eficacia, y usa las distintas alternativas de diseño experimental donde algunos sesgos estadísticos han sido neutralizados.

- El **Capítulo 8** reformula el problema de recomendación basado en usuarios, dando una generalización del mismo como un problema de predicción de eficacia. Investiga cómo adoptar dicha generalización para definir un marco unificado donde podamos realizar un análisis objetivo de la efectividad (poder predictivo) de las funciones de valoración de los vecinos.

Parte V. Conclusiones

- El **Capítulo 9** concluye con un resumen de las principales contribuciones de esta tesis y una discusión sobre las líneas de investigación futuras.

Parte VI. Apéndices

- El **Apéndice A** da detalles sobre los métodos propuestos en esta tesis: configuración de los algoritmos de recomendación y los parámetros de los diseños experimentales usados en la evaluación. También reporta estadísticas detalladas sobre los conjuntos de datos utilizados en los experimentos, complementando otros datos incluidos en capítulos previos.
- El **Apéndice B** contiene la traducción a español del Capítulo 1.
- El **Apéndice C** contiene la traducción a español del Capítulo 9.