

Sistema de Recomendación Híbrido

Marian Susana Álvarez Suri

Carlos Antonio Bresó Sotto

<https://github.com/marians002/IRS-Final-Project.git>

Facultad de Matemática y Computación, Universidad de La Habana, Cuba

Resumen. Este trabajo presenta un sistema de recomendación híbrido que integra métodos basados en contenido y filtrado colaborativo para mejorar la precisión y relevancia de las recomendaciones. El sistema propuesto utiliza un modelo de filtrado colaborativo para capturar las preferencias de los usuarios a partir de sus interacciones pasadas, mientras que el enfoque basado en contenido analiza las características intrínsecas de los ítems para ofrecer recomendaciones personalizadas. La combinación de estos métodos se realiza mediante un esquema de ponderación que ajusta la contribución de cada técnica según el contexto y el perfil del usuario. Los resultados experimentales muestran que el sistema híbrido supera a los enfoques individuales en términos de precisión y diversidad.

Palabras clave: Sistema de Recomendación Híbrido · Filtrado colaborativo · Filtrado basado en contenido · Recomendación de películas.

1. Introducción

Los sistemas de recomendación han emergido como uno de los pilares fundamentales en la estrategia digital de las grandes compañías a nivel mundial, transformándose en herramientas indispensables para mejorar la experiencia del usuario y maximizar la eficiencia de los servicios en línea.

El proyecto en cuestión aborda la creación de un sistema de recomendación híbrido para películas. Para ello se combinan las fortalezas del filtrado colaborativo y del filtrado basado en contenido con el objetivo de superar las limitaciones inherentes a cada uno de estos enfoques por separado. El filtrado colaborativo aprovecha las interacciones y preferencias de los usuarios para hacer recomendaciones, basándose en la premisa de que aquellos con gustos similares en el pasado tendrán intereses parecidos en el futuro. Por otro lado, el filtrado basado en contenido se centra en las características de los ítems para hacer recomendaciones, lo que permite una personalización más detallada, pero puede limitar la diversidad de las recomendaciones.

El objetivo general de este trabajo es desarrollar un sistema que no solo se base en las preferencias explícitas del usuario, sino que también incorpore elementos que podrían resultar atractivos a pesar de su diferencia con el historial previo o con los gustos de personas con intereses similares. Esta estrategia busca enriquecer la diversidad de las recomendaciones cinematográficas, permitiendo descubrir nuevas películas y opciones que de otro modo podrían pasar desapercibidas.

2. Estado del Arte

Los sistemas de recomendación se han vuelto indispensables en plataformas digitales como Netflix, Amazon y Spotify, donde la personalización y la relevancia de las recomendaciones son clave para la satisfacción del usuario. En este contexto, los sistemas de recomendación han evolucionado significativamente, pasando de enfoques basados en filtrado colaborativo y contenido a sistemas híbridos que combinan múltiples técnicas para mejorar la precisión y diversidad de las recomendaciones [1]. Entre las técnicas más destacadas para abordar este problema, encontramos el filtrado colaborativo, que se basa en la similitud entre usuarios o ítems para hacer recomendaciones, y el filtrado basado en contenido, que analiza las características intrínsecas de los ítems para generar recomendaciones personalizadas. Ambos enfoques han demostrado ser efectivos en diferentes contextos, pero presentan limitaciones en términos de precisión, diversidad y escalabilidad [2].

2.1. Actualidad en Sistemas Híbridos

En el ámbito de los sistemas de recomendación híbridos, el estado del arte ha evolucionado significativamente, reflejando un creciente interés en superar las limitaciones de los enfoques tradicionales mediante la integración de múltiples métodos de recomendación. Según el análisis presentado en el artículo "Hybrid recommender systems: A systematic literature review"[3], los sistemas de recomendación híbridos ofrecen una solución efectiva para superar las limitaciones de los sistemas basados únicamente en filtrado colaborativo o contenido, pues han demostrado ser especialmente eficaces en abordar problemas como el arranque frío y la escasez de datos, ofreciendo recomendaciones más precisas y personalizadas. Los estudios [4] destacan la diversidad de técnicas utilizadas para combinar diferentes enfoques, desde métodos ponderados hasta la fusión de características, mostrando una tendencia hacia la experimentación y la innovación en el campo. Existen varias formas de combinar estos métodos [2], entre las cuales destacan:

1. **Ponderación de Pesos:** En este enfoque, las recomendaciones de diferentes sistemas se combinan asignando un peso a cada una de ellas. La importancia o puntuación de cada producto se calcula a partir de los resultados obtenidos por todas las técnicas de recomendación presentes en el sistema. Un ejemplo de este tipo de hibridación es el sistema P-Tango [4], que ajusta los pesos de acuerdo con las opiniones de los usuarios sobre los resultados mostrados.
2. **Mezcla de Modelos:** En este método, las recomendaciones provenientes de varias técnicas de recomendación se muestran al mismo tiempo [2]. Esto permite que el usuario vea una combinación de recomendaciones basadas en diferentes enfoques, lo que puede mejorar la diversidad y la calidad de las recomendaciones.
3. **Sistema Híbrido Alternante:** Este enfoque utiliza criterios específicos para alternar entre diferentes técnicas de recomendación. Por ejemplo, primero se emplean técnicas basadas en contenido y, si no se obtienen buenos

resultados, el sistema cambia a filtrado colaborativo. Esto permite que las recomendaciones sean sensibles a las fortalezas y debilidades de cada método [5].

4. **Combinación de Características:** En este método, las características de las fuentes de datos de varias técnicas de recomendación se combinan en un único algoritmo de recomendación [2]. Esto permite aprovechar la información de diferentes fuentes para generar recomendaciones más precisas.

Las revisiones sistemáticas [3] [6] subrayan la importancia de los sistemas de recomendación híbridos en el panorama actual, identificando áreas clave para la investigación futura, como la mejora de los algoritmos de fusión y la exploración de nuevas fuentes de datos para enriquecer las recomendaciones [1]. La continua evolución de estos sistemas promete no solo mejorar la experiencia del usuario en plataformas digitales, sino también abrir nuevas vías para el avance tecnológico en la inteligencia artificial y el análisis de datos [6].

3. Implementación

El sistema desarrollado es un recomendador híbrido que combina múltiples algoritmos de recomendación mediante la técnica de ponderación de pesos. Este sistema se encarga de generar recomendaciones personalizadas para los usuarios basándose en sus interacciones previas con los ítems, en la similitud con otros usuarios, y en la diversidad de géneros.

Además, se mejoran las sugerencias finales incorporando opciones derivadas de películas con géneros distintos a los inicialmente seleccionados, enriqueciendo así las propuestas presentadas al usuario.

3.1. Datos Utilizados

Para generar las recomendaciones se utiliza el conjunto de datos MovieLens 100K, que contiene 100,000 calificaciones de películas realizadas por 943 usuarios sobre 1682 películas. Este conjunto de datos incluye información sobre las películas, como títulos, géneros y fechas de lanzamiento, así como las calificaciones otorgadas por los usuarios [7].

3.2. Algoritmos Empleados

Se tomaron los siguientes algoritmos de recomendación de la librería *cornac* [8] para implementar el sistema híbrido:

- SVD (Singular Value Decomposition): Un modelo de factorización de matrices que descompone la matriz de calificaciones en matrices de características latentes [9].
- ItemKNN (Item-based K-Nearest Neighbors): Un modelo basado en la similitud entre ítems, utilizando la similitud coseno [8].
- BPR (Bayesian Personalized Ranking): Un modelo de optimización para el ranking personalizado basado en la probabilidad bayesiana [10].

3.3. Integración de Técnicas

Para combinar los modelos mencionados utilizamos un esquema de ponderación. Los pesos asignados a cada modelo son (4, 1, 6) para SVD, BPR e ItemKNN, respectivamente. Estos pesos fueron elegidos porque recompensan positivamente a los algoritmos que mejor desempeño tuvieron individualmente al recomendar ítems novedosos.

Además, se calcula la similitud entre las películas basándose en sus géneros utilizando la similitud de coseno. Esta información de similitud se utiliza para mejorar las recomendaciones, añadiendo películas no tan similares a las mejor recomendadas.

4. Evaluación del Sistema

Para evaluar el rendimiento del sistema de recomendación híbrido, se utilizan varias métricas estándar en el campo de los sistemas de recomendación [11]. Estas métricas proporcionan una evaluación cuantitativa de la precisión, relevancia y diversidad de las recomendaciones generadas por el sistema, permitiendo una comparación objetiva con otros enfoques y configuraciones. Las métricas empleadas se explican a continuación:

1. **Mean Absolute Error (MAE)**: Mide el error promedio absoluto entre las calificaciones predichas y las calificaciones reales. Un MAE bajo indica que las predicciones del sistema están cercanas a las valoraciones reales. Se elige por su interpretabilidad directa en términos de error promedio en las predicciones [12].
2. **Root Mean Square Error (RMSE)**: Es la raíz cuadrada del error cuadrático medio. Penaliza más los errores grandes que el MAE, lo que significa que un RMSE alto indica la presencia de errores grandes en algunas predicciones. Se selecciona por su sensibilidad a los errores grandes, lo que es crucial en sistemas de recomendación donde es importante minimizar las malas recomendaciones [12].
3. **Precision@10**: Precision@10 mide la proporción de recomendaciones relevantes entre las top-10 recomendaciones hechas por el sistema. [13] Una Precision@10 alta indica que la mayoría de las top-10 recomendaciones son relevantes para el usuario. Se elige para evaluar la calidad y relevancia de las recomendaciones en la parte superior de la lista, lo cual es importante desde la perspectiva del usuario, ya que los usuarios suelen prestar atención solo a las primeras recomendaciones.
4. **Recall@10**: Recall@10 mide cuántos ítems relevantes se encuentran entre las top-10 recomendaciones del sistema, respecto al total de ítems relevantes. Un Recall@10 alto indica que el sistema es capaz de identificar una gran proporción de los ítems relevantes en sus top-10 recomendaciones. Se utiliza para asegurar que el sistema no solo haga recomendaciones precisas sino que también capture una buena parte de los ítems relevantes en sus recomendaciones principales [13].

5. **Novelty**: Permite medir cuán desconocidos o nuevos son los ítems recomendados para el usuario [14]. Se puede adaptar para evaluar la diversidad al considerar cómo de variadas son las recomendaciones entre sí. La inclusión de la novedad para evaluar la diversidad asegura que el sistema no solo recomiende ítems relevantes, sino que también promueva el descubrimiento y la exploración al sugerir ítems variados y potencialmente inesperados.

5. Resultados Obtenidos

En esta sección, presentamos una comparativa exhaustiva de las métricas de evaluación obtenidas por los modelos empleados. La Tabla 1 resume los resultados en términos de Error Medio Absoluto (MAE), Raíz del Error Cuadrático Medio (RMSE), Precisión y Recall considerando los primeros 10 ítems recomendados (Precision@10 y Recall@10). Estos resultados nos permiten analizar la eficacia y eficiencia de nuestro sistema híbrido, comparándolo con métodos tradicionales como Matrix Factorization (MF), Bayesian Personalized Ranking (BPR), Singular Value Decomposition (SDV) y K-Nearest Neighbors for Items (ItemKNN).

Cuadro 1. Evaluación combinada del sistema de recomendación

Modelo	MAE	RMSE	Precision@10	Recall@10	Novedad
MF	0.7435	0.9023	0.0672	0.0464	0.02052
BPR	2.1572	2.3603	0.1138	0.1120	0.00254
SVD	0.7494	0.9084	0.0574	0.0397	0.02707
ItemKNN	0.8222	0.9944	0.0305	0.0176	0.26961
Hybrid	0.7968	0.9476	0.1356	0.1042	0.32069

5.1. Análisis de los resultados

Al analizar los resultados de la tabla podemos observar que el sistema MF es el más preciso en términos de error de predicción, con los valores más bajos de MAE y RMSE. Sin embargo, el sistema híbrido destaca en Precision@10 con un valor de 0.1356, indicando que tiene la mayor proporción de recomendaciones relevantes entre las top-10 recomendaciones. En términos de Recall@10, el modelo BPR lidera con 0.1120, seguido de cerca por el sistema híbrido con 0.1042, lo que sugiere que ambos modelos son efectivos en capturar una amplia gama de ítems relevantes. Además, el sistema híbrido muestra la mayor puntuación de novedad con 0.32069, lo que indica su capacidad para recomendar ítems novedosos y diversos. En contraste, los modelos SVD y ItemKNN presentan limitaciones tanto en precisión como en cobertura de ítems relevantes, con valores más bajos en Precision@10 y Recall@10.

6. Limitaciones

A pesar de las ventajas del sistema híbrido implementado, existen algunas limitaciones que deben considerarse. En primer lugar, aunque se ha implementado una funcionalidad para aumentar la diversidad añadiendo recomendaciones de géneros diferentes a los más sugeridos, esto podría llevar a recomendaciones que no sean del todo relevantes para el usuario. La inclusión de ítems de géneros menos preferidos puede mejorar la diversidad, pero también puede disminuir la precisión y la satisfacción del usuario si estos ítems no coinciden con sus intereses.

Otra limitación es la complejidad computacional y el tiempo de procesamiento. La combinación de múltiples modelos y la ponderación de sus resultados pueden aumentar significativamente el tiempo de entrenamiento y prueba, lo cual podría representar un desafío en entornos donde se requiera una respuesta rápida y eficiente.

7. Conclusiones

El sistema híbrido, aunque no es el más preciso en términos de predicción de calificaciones, demuestra ser el más efectivo en ofrecer recomendaciones relevantes. Además, su desempeño sugiere que es capaz de capturar una amplia gama de ítems relevantes, lo que podría indicar una mayor diversidad en sus recomendaciones. Esto lo hace particularmente valioso en escenarios donde la relevancia y la diversidad de las recomendaciones son críticas para la experiencia del usuario. Por tanto, el sistema híbrido ofrece un equilibrio óptimo entre precisión, cobertura y diversidad, superando a los métodos tradicionales en varios aspectos clave.

Referencias

1. G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE transactions on knowledge and data engineering*, vol. 17, no. 6, pp. 734–749, 2005.
2. R. Burke, "Hybrid recommender systems: Survey and experiments," *User modeling and user-adapted interaction*, vol. 12, pp. 331–370, 2002.
3. E. Çano and M. Morisio, "Hybrid recommender systems: A systematic literature review," *Intelligent Data Analysis*, vol. 21, no. 6, pp. 1487–1524, 2017.
4. L. G. P. Córdón, "Modelos de recomendación con falta de información: aplicaciones al sector turístico," *Consejería de Turismo, Comercio y Deporte, Junta de Andalucía*, vol. 1, no. N/A, p. N/A, 2009, tesis doctoral dirigida por Luis Martínez López, Universidad de Jaén, Escuela Politécnica Superior, Departamento de Informática.
5. C. V. Sacín, "Propuesta de una metodología de aplicación de técnicas de descubrimiento del conocimiento para la ayuda al estudiante en entornos de enseñanza superior," *Universidad Autónoma de Madrid, Escuela Politécnica Superior, Departamento de Ingeniería Informática*, vol. N/A, no. N/A, p. N/A, 2010, tesis doctoral dirigida por D. Álvaro Ortigosa.

6. D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *Journal of Big Data*, vol. 9, no. 1, p. Article number: 59, 2022.
7. "Movielens dataset," (<https://grouplens.org/datasets/movielens/>), 2024.
8. H. W. L. Masaaki Kotera, "Cornac: A comparative framework for multimodal recommender systems," *Journal of Machine Learning Research*, vol. 21, no. 95, pp. 1–5, 2020. [Online]. Available: <http://jmlr.org/papers/v21/19-1015.html>
9. C. V. Yehuda Koren, Robert Bell, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.
10. S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*, 2009, pp. 452–461.
11. L. Monigatti, "Evaluation metrics for search and recommendation systems," <https://weaviate.io/blog/retrieval-evaluation-metrics>, May 2024.
12. C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance," *Climate Research*, vol. 30, no. 1, pp. 79–82, 2005.
13. D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011. [Online]. Available: https://www.researchgate.net/publication/220708622_Evaluation_From_Precision_Recall_and_F-Measure_to_ROC_Informedness_Markedness_and_Correlation
14. S. Vargas and P. Castells, "Rank and relevance in novelty and diversity metrics for recommender systems," *Proceedings of the Fifth ACM Conference on Recommender Systems*, pp. 109–116, 2011.