

Sistemas de Recomendación

Víctor González

Contenido

- Introducción
- Definición
- Tipos de sistemas de recomendación
- Algoritmos comunes utilizados en sistemas de recomendación
- Desarrollo del Sistema de Recomendación Académica
- Evaluación del Sistema de Recomendación
- Consideraciones lingüísticas
- Consideraciones Éticas
- Limitante de los sistemas de recomendación
- Conclusiones

Introducción

- Los sistemas de recomendación son herramientas poderosas que utilizan algoritmos para predecir y sugerir elementos de interés para los usuarios. En el contexto empresarial, estos sistemas juegan un papel fundamental al proporcionar recomendaciones personalizadas de productos o servicios a los clientes, lo que aumenta la satisfacción del cliente y las ventas. Al analizar el comportamiento pasado del usuario, así como sus preferencias y patrones de compra, los sistemas de recomendación pueden generar sugerencias relevantes y adaptadas a cada individuo, mejorando así la experiencia del usuario y fomentando la fidelidad a la marca. Estos sistemas ayudan a las empresas a aumentar las ventas, comprender mejor el comportamiento del consumidor y a tomar decisiones más informadas sobre estrategias de marketing y desarrollo de productos.
- En adición, estas herramientas ayudan a los usuarios a descubrir nuevos productos o servicios que podrían ser de su interés, ahorrándoles tiempo y esfuerzo en la búsqueda manual. Además, al recibir recomendaciones personalizadas, los usuarios pueden sentirse más valorados y comprendidos por la empresa, lo que fortalece la relación cliente-empresa.
- Los sistemas de recomendación son una pieza clave en el ecosistema empresarial moderno, proporcionando beneficios tanto para las empresas como para los consumidores al facilitar la toma de decisiones y mejorar la experiencia de compra.

Definición

- Los sistemas de recomendación son sistemas de software diseñados para recomendar elementos relevantes a usuarios, basándose en su perfil o preferencias. Estos sistemas son ampliamente utilizados para ayudar a los usuarios a descubrir nuevos elementos de interés, como libros, películas, artículos científicos o cursos, entre otros. La recomendación se realiza generalmente mediante el análisis de datos históricos de usuarios y elementos, utilizando algoritmos que predicen la afinidad entre usuarios y elementos.

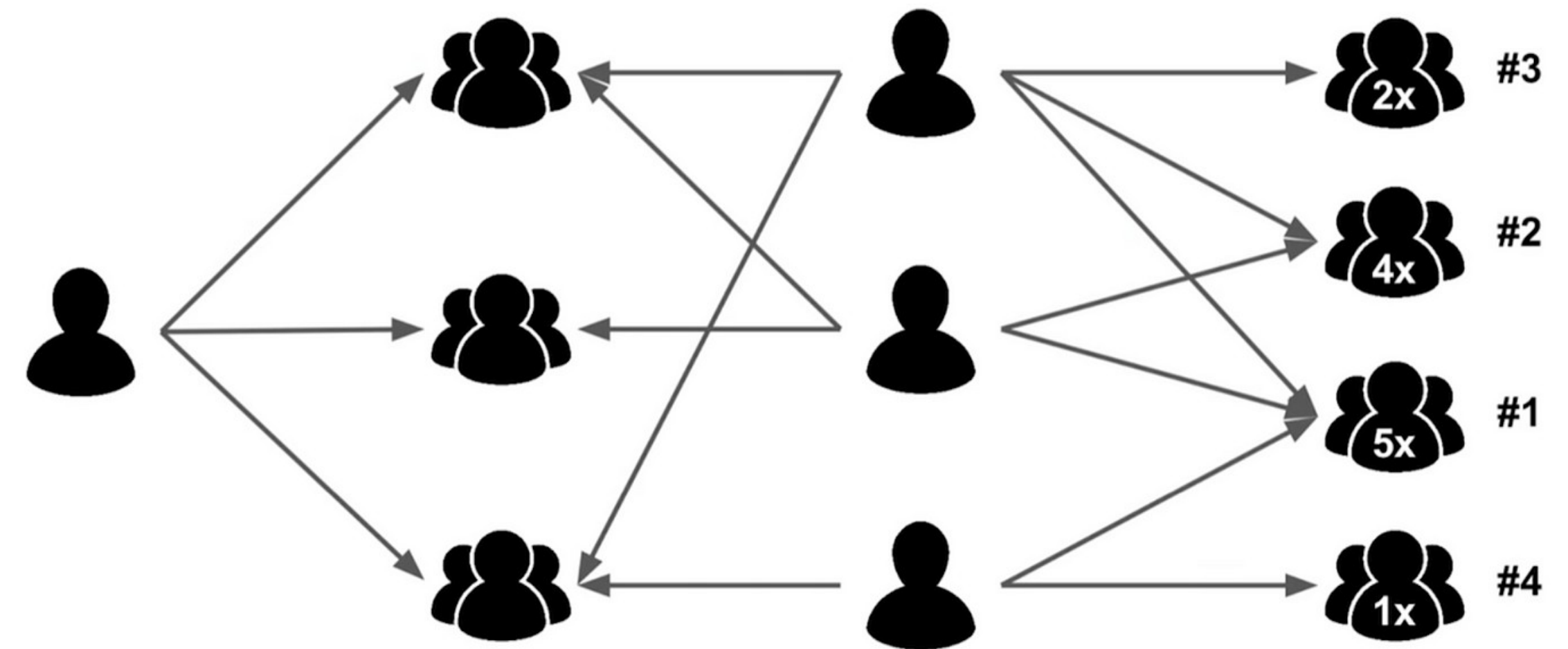
Tipos de sistemas de recomendación

Filtrado colaborativo

El filtrado colaborativo es ampliamente utilizado en sistemas de recomendación debido a su capacidad para generar recomendaciones personalizadas al aprovechar la información de las interacciones de usuarios.

- **Basado en Usuarios:** Este enfoque compara las preferencias de un usuario con las de otros usuarios similares. Se utiliza la información de las interacciones pasadas de un usuario y se encuentran usuarios con perfiles similares cuyas preferencias se pueden utilizar para hacer recomendaciones.
- **Basado en Ítems:** En lugar de comparar usuarios entre sí, este método se centra en comparar los ítems entre sí. Se analiza la similitud entre los ítems según cómo han sido evaluados o consumidos por los usuarios en el pasado. Luego, se recomiendan ítems similares a los que el usuario ha mostrado interés previamente.
- **Modelos de Factorización:** Estos modelos representan a usuarios e ítems en un espacio de características latentes. Utilizan técnicas de factorización matricial para encontrar patrones en las interacciones usuario-ítem y generar recomendaciones. Este enfoque es particularmente útil cuando hay datos escasos o ruidosos.
- **Filtrado Colaborativo Basado en Modelos:** Combina métodos de filtrado colaborativo con técnicas de aprendizaje automático para mejorar la precisión de las recomendaciones. Estos modelos pueden capturar relaciones más complejas entre usuarios e ítems y proporcionar recomendaciones más personalizadas.

User-based collaborative filtering algorithm

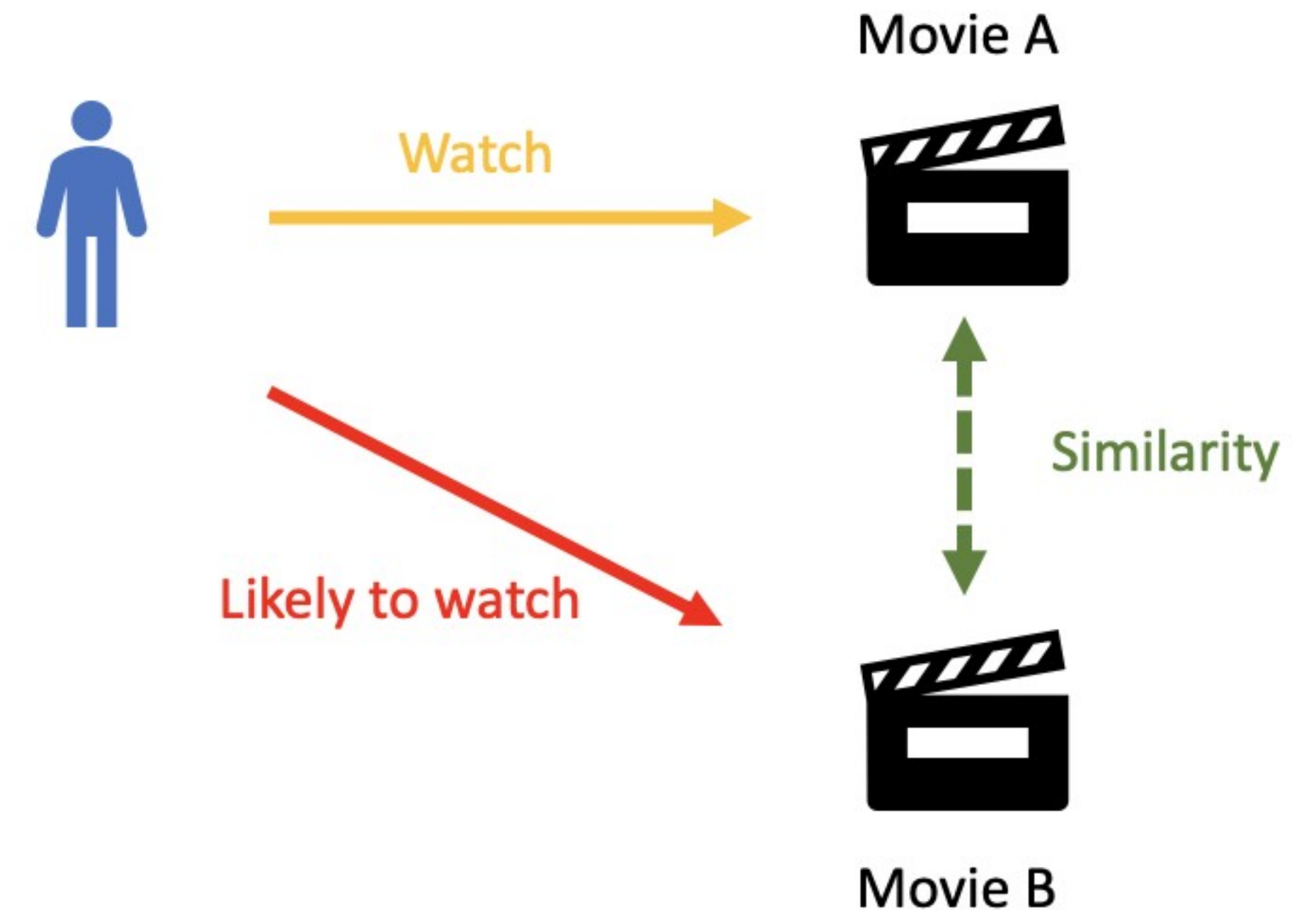


Tipos de sistemas de recomendación

Filtrado basado en contenido

El filtrado basado en contenido es efectivo para recomendar elementos similares a los que un usuario ha interactuado en el pasado, ya que se basa en el análisis de las características de los elementos y las preferencias del usuario.

- **Análisis de Características:** Este enfoque analiza las características de los elementos para comprender su naturaleza. Por ejemplo, en el caso de películas, las características podrían incluir el género, el director, el reparto, la trama, etc. Estas características se utilizan luego para encontrar elementos similares.
- **Perfil de Usuario:** Se construye un perfil de usuario basado en las preferencias o historial de interacciones del usuario con los elementos. Luego, se buscan elementos cuyas características se correlacionen con las preferencias del usuario para hacer recomendaciones personalizadas.
- **Modelos de Aprendizaje Automático:** Al igual que en el filtrado colaborativo, se pueden utilizar técnicas de aprendizaje automático para mejorar la precisión de las recomendaciones. Estos modelos pueden aprender patrones complejos en los datos y proporcionar recomendaciones más precisas y personalizadas.
- **Filtrado Basado en Contenido Multimodal:** Además de analizar características textuales, este enfoque también considera otros tipos de datos, como imágenes, audio o video. Por ejemplo, en un sistema de recomendación de películas, además de las características textuales como el género y la trama, también se podrían considerar características visuales como el cartel de la película o características auditivas como la banda sonora.



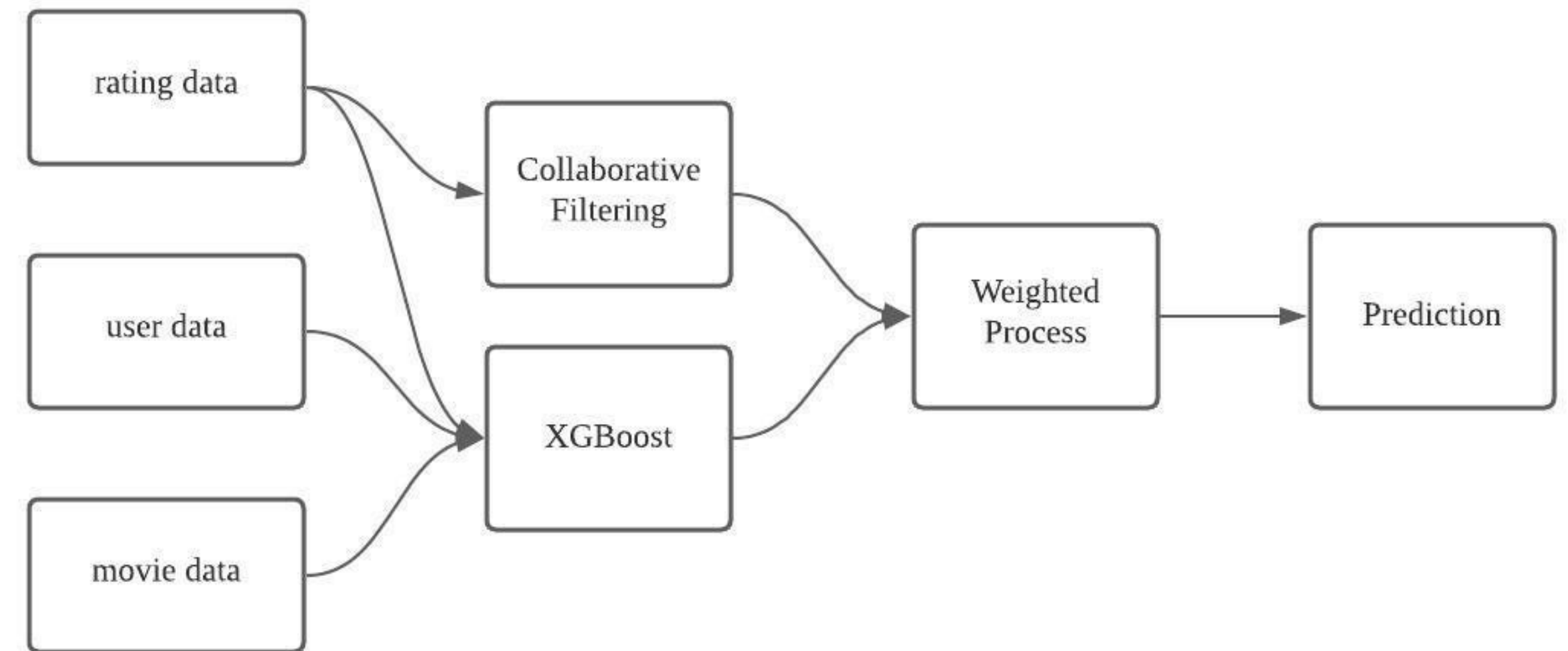
Leban, J. (2021, December 14). Essentials of recommendation engines: content-based and collaborative filtering. *Medium*. <https://towardsdatascience.com/essentials-of-recommendation-engines-content-based-and-collaborative-filtering-31521c964922>

Tipos de sistemas de recomendación

Sistemas híbridos

Los sistemas híbridos son ampliamente utilizados en sistemas de recomendación debido a su capacidad para mejorar la calidad y la diversidad de las recomendaciones al combinar múltiples enfoques. Este enfoque permite crear sistemas más robustos y adaptativos que pueden satisfacer mejor las necesidades y preferencias de los usuarios.

- **Combinación de Enfoques:** Los sistemas híbridos combinan múltiples enfoques de recomendación, como el filtrado colaborativo y el filtrado basado en contenido, para mejorar la calidad de las recomendaciones. Esto permite aprovechar las fortalezas de cada enfoque y compensar sus debilidades.
- **Modelos de Fusión:** En estos sistemas, se utilizan modelos de fusión para combinar las recomendaciones generadas por diferentes métodos en una única lista de recomendaciones. Esto puede hacerse de manera ponderada, dando más peso a ciertos métodos dependiendo del contexto o la situación.
- **Adaptación Dinámica:** Algunos sistemas híbridos pueden adaptarse dinámicamente según el contexto o las preferencias del usuario. Por ejemplo, podrían cambiar entre diferentes enfoques de recomendación dependiendo del tipo de elemento o del historial de interacciones del usuario.
- **Personalización:** Estos sistemas pueden ofrecer recomendaciones personalizadas al considerar las preferencias individuales de los usuarios y adaptar las recomendaciones en función de su perfil y comportamiento.
- **Reducción de Ruido y Esparsidad:** Al combinar diferentes fuentes de información, los sistemas híbridos pueden reducir el ruido y la esparsidad en los datos, lo que lleva a recomendaciones más precisas y relevantes.

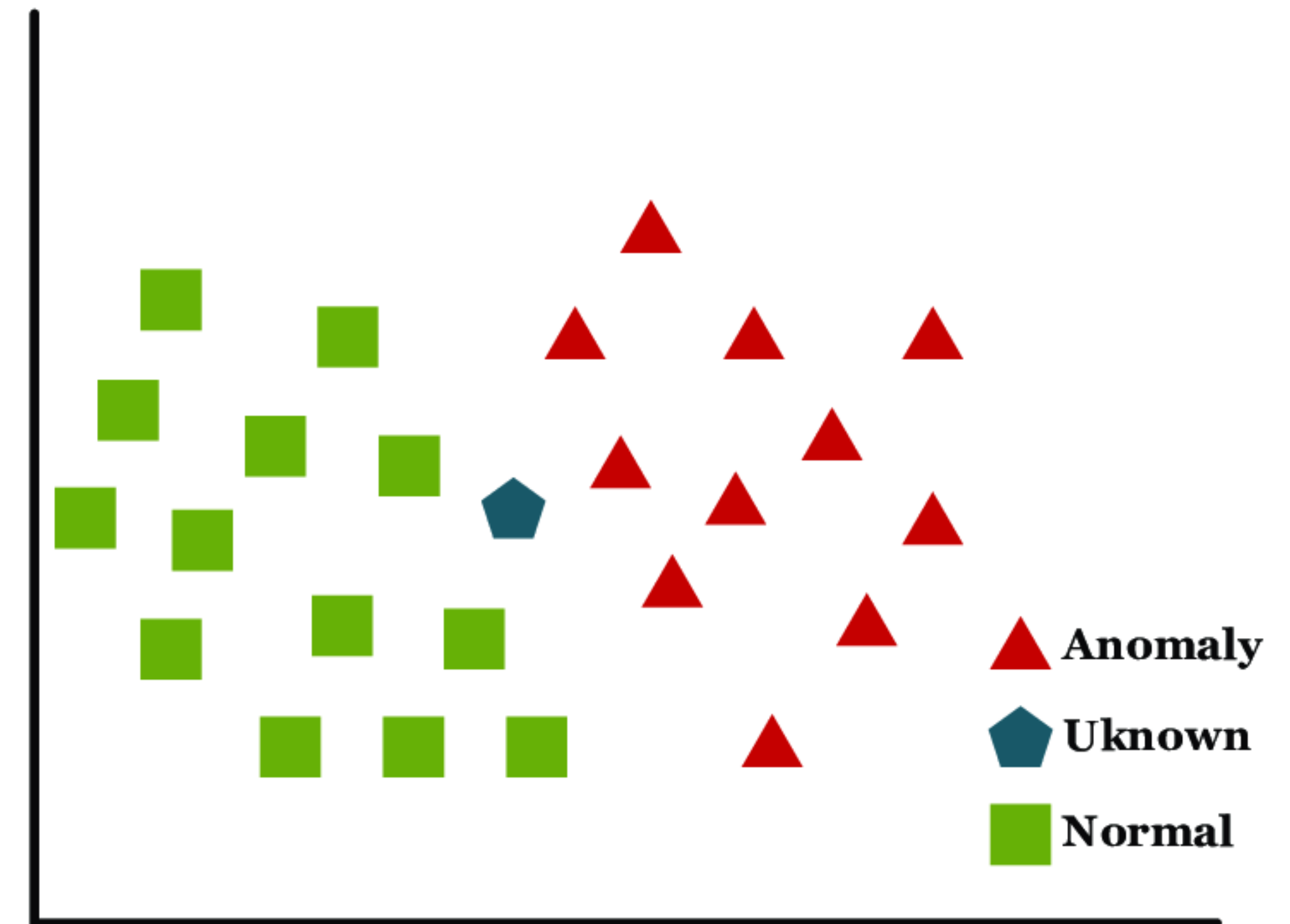


Chiang, J. (2024, March 14). 7 Types of hybrid recommendation System - Analytics Vidhya - Medium. *Medium*.
<https://medium.com/analytics-vidhya/7-types-of-hybrid-recommendation-system-3e4f78266ad8>

Algoritmos comunes

Filtrado colaborativo

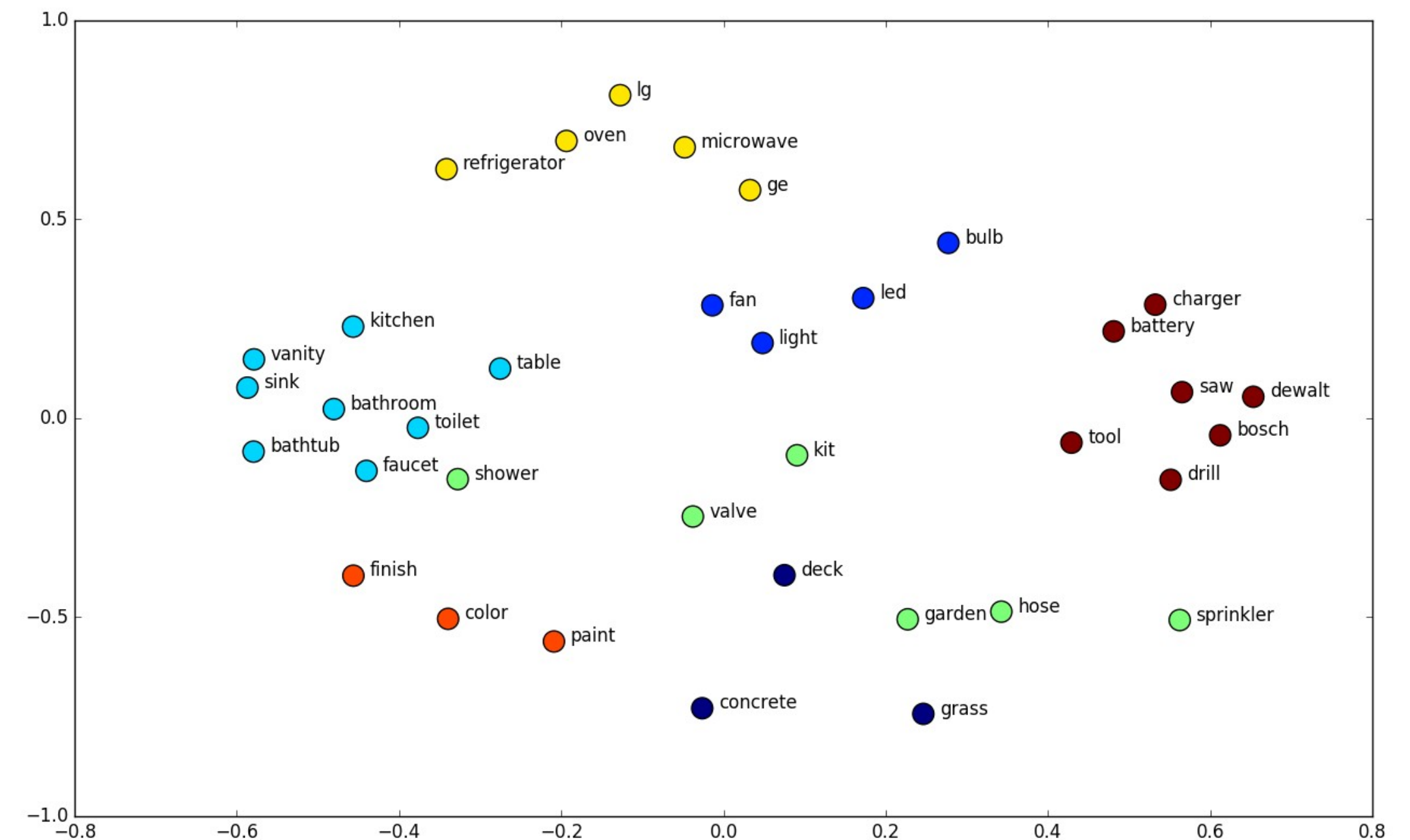
- Basado en Memoria: Este enfoque utiliza la similitud entre usuarios o ítems para hacer recomendaciones. Algunos de los métodos más comunes son el filtrado colaborativo basado en vecinos más cercanos (k-NN) y el filtrado colaborativo basado en descomposición de valores singulares (SVD).
- Basado en Modelos: Estos métodos utilizan técnicas de aprendizaje automático, como modelos de factorización matricial, para aprender patrones en los datos de interacción usuario-ítem y generar recomendaciones.



Villalonga, R. (2021, August 9). Filtrado Colaborativo (FC) basado en el usuario - Rodrigo Villalonga - Medium. *Medium*.
<https://medium.com/@rvillalongar/filtrado-colaborativo-fc-basado-en-el-usuario-f13926ffdfac>

Algoritmos comunes Filtrado basado en contenido

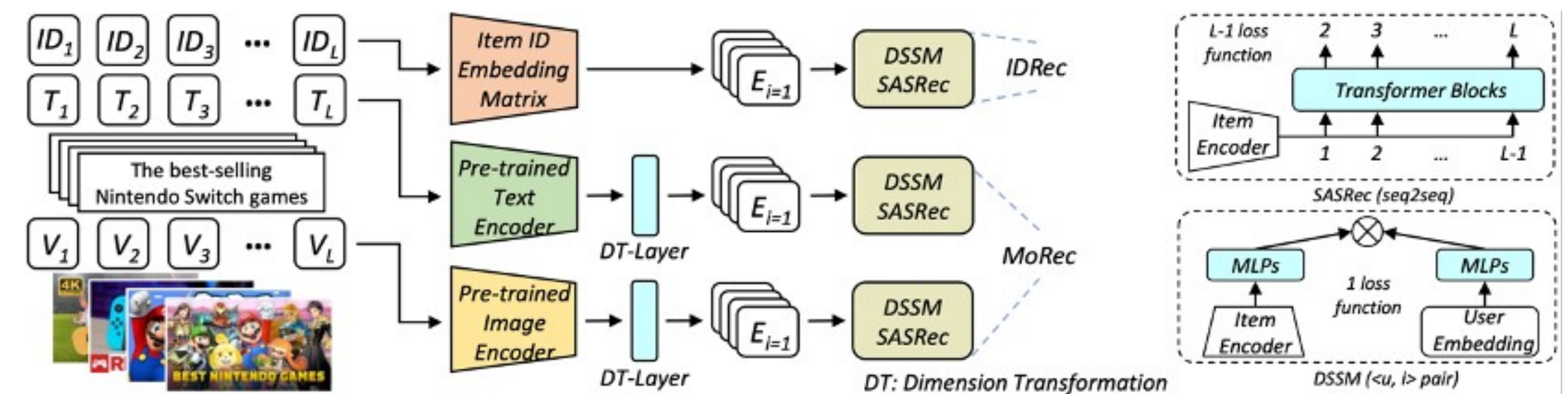
- TF-IDF: Este método asigna un peso a cada término en función de su frecuencia en un documento y de su rareza en el corpus. Luego, se calcula la similitud entre documentos utilizando estas ponderaciones.
- Vectores de Palabras (Word Embeddings): Utilizando técnicas como Word2Vec o GloVe, se representan las palabras en un espacio vectorial donde palabras similares tienen vectores cercanos. Esto se utiliza para calcular la similitud entre elementos basados en sus descripciones o características.



Barla, N. (2023, August 18). *The Ultimate Guide to word embeddings*. neptune.ai. <https://neptune.ai/blog/word-embeddings-guide>

Algoritmos comunes Sistemas Híbridos

- Fusión de Modelos: Se combinan múltiples modelos de recomendación, ya sea basados en colaboración, contenido u otros, utilizando técnicas de fusión como el promedio, la ponderación o el voto mayoritario.
- Conjuntos de Modelos: Se utilizan conjuntos de modelos para generar múltiples recomendaciones y luego combinarlas en una lista final. Esto puede incluir métodos como el bagging, boosting o ensambles de modelos.

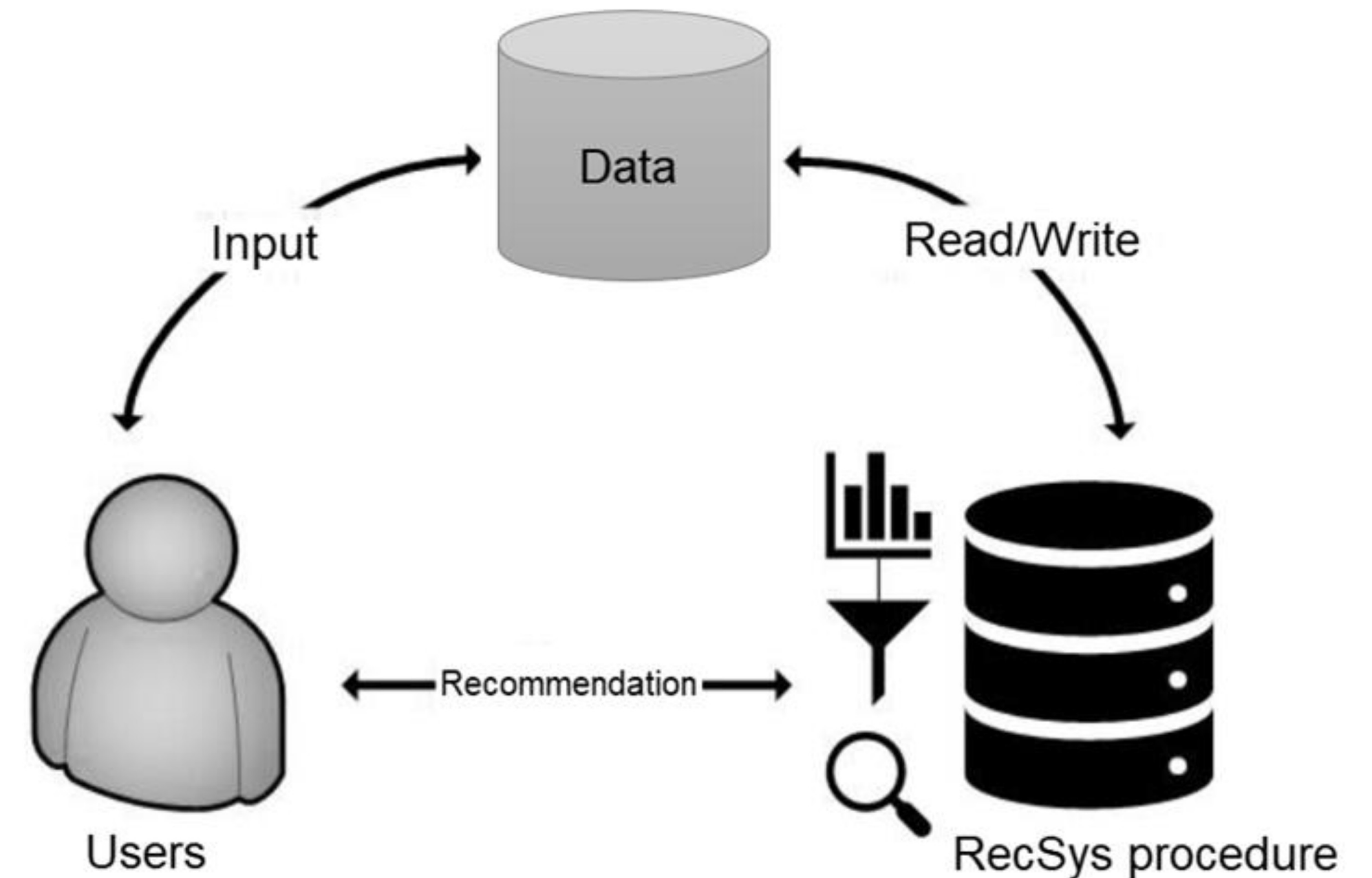


Ai-Advance. (2024, January 4). Multimodal Recommender System vs. ID-based Recommender System Revisited. *Medium*. https://medium.com/@lifengyi_6964/multimodal-recommender-system-vs-id-based-recommender-system-revisited-588ca88cd16e

Desarrollo del Sistema de Recomendación Académica

Recolección de Datos

- Identificación de Fuentes: Determinar las fuentes de datos relevantes para el sistema de recomendación, como bases de datos académicas, repositorios institucionales, registros de usuarios, etc.
- Obtención de Datos: Recopilar datos relacionados con la actividad académica, como publicaciones científicas, historial de cursos, intereses de investigación, calificaciones, etc.

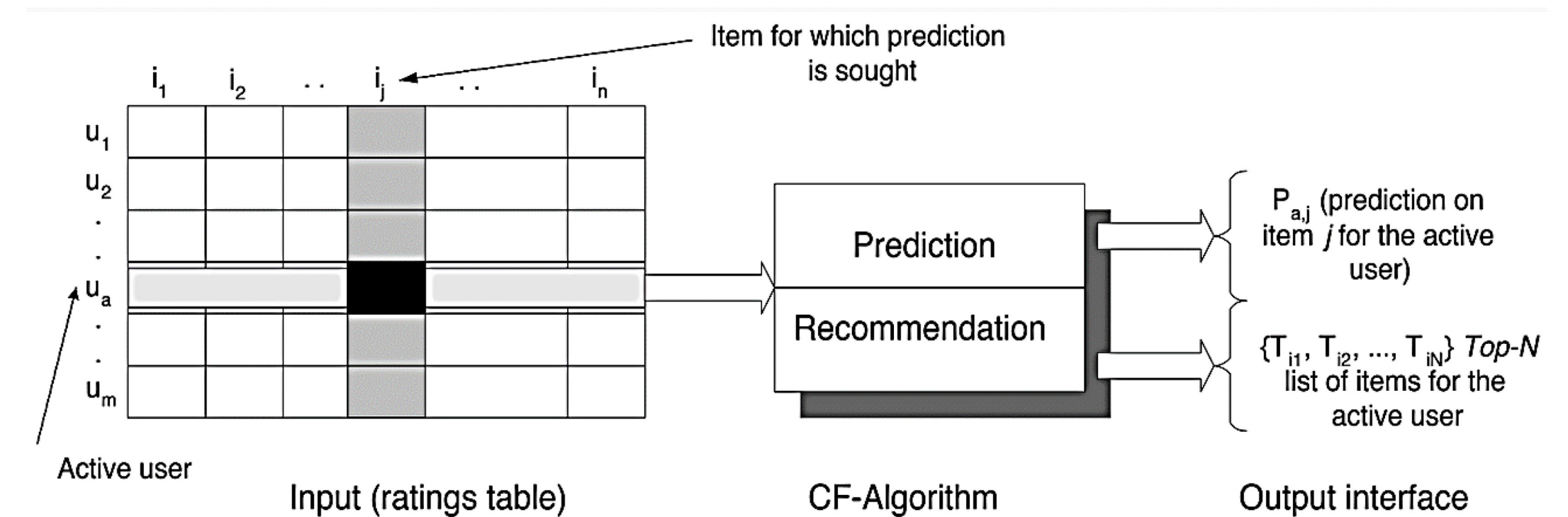


Azeroual, O., & Koltay, T. (2022). RecSys Pertaining to Research Information with Collaborative Filtering Methods: Characteristics and Challenges. *Publications*, 10(2), 17. <https://doi.org/10.3390/publications10020017>

Desarrollo del Sistema de Recomendación Académica

Preprocesamiento de Datos

- Limpieza de Datos: Eliminar datos redundantes, incompletos o incorrectos que puedan afectar la calidad de las recomendaciones.
- Integración de Datos: Integrar datos de múltiples fuentes en un formato coherente y compatible para su posterior análisis.
- Transformación de Datos: Convertir los datos en un formato adecuado para su procesamiento y análisis, como matrices de usuario-ítem o matrices de características.



Azeroual, O., & Koltay, T. (2022). RecSys Pertaining to Research Information with Collaborative Filtering Methods: Characteristics and Challenges. *Publications*, 10(2), 17. <https://doi.org/10.3390/publications10020017>

Desarrollo del Sistema de Recomendación Académica

Representación de Datos

- Extracción de Características: Identificar las características relevantes de los datos, como palabras clave en publicaciones académicas o temas de interés del usuario.
- Vectorización: Representar los datos en forma de vectores numéricos para que puedan ser procesados por algoritmos de aprendizaje automático, utilizando técnicas como TF-IDF, Word Embeddings, entre otras.

	i_1	i_2	...	i_j	...	i_k	...	i_m
u_1				R		R		
u_2				R		R		
\vdots				\vdots		\vdots		
u_i				R		R		
\vdots				\vdots		\vdots		
u_n				R		R		

Desarrollo del Sistema de Recomendación Académica

Modelado de Recomendación

- Selección de Algoritmos: Elegir los algoritmos de recomendación adecuados según el tipo de datos y el contexto de aplicación, como filtrado colaborativo, filtrado basado en contenido o modelos híbridos.
- Entrenamiento del Modelo: Entrenar el modelo de recomendación utilizando datos históricos de interacciones usuario-ítem para aprender patrones y generar recomendaciones.
- Evaluación del Modelo: Evaluar el rendimiento del modelo utilizando métricas como precisión, recall, RMSE, etc., para garantizar su eficacia y calidad.

	p'	n'
p	TP = recommended and read	FN = recommended but not read
n	FP = not recommended but read	TN = not recommended but not read

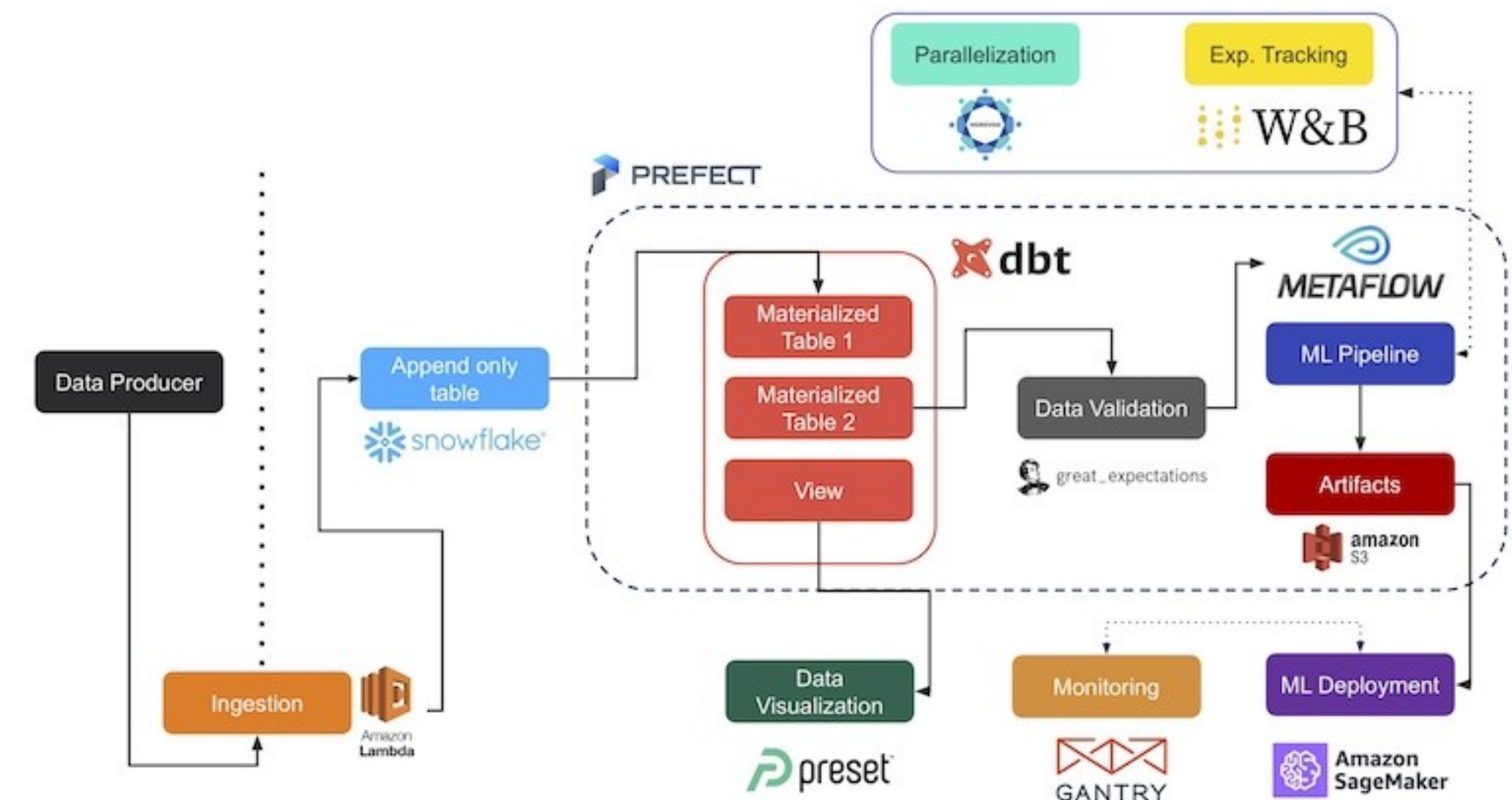
p = positive
n = negative
TP = true positive
FP = false positive
TN = true negative
FN = false negative

Azeroual, O., & Koltay, T. (2022). RecSys Pertaining to Research Information with Collaborative Filtering Methods: Characteristics and Challenges. *Publications*, 10(2), 17. <https://doi.org/10.3390/publications10020017>

Desarrollo del Sistema de Recomendación Académica

Implementación del Sistema

- Desarrollo de la Interfaz: Diseñar e implementar una interfaz de usuario intuitiva y fácil de usar para que los usuarios puedan interactuar con el sistema de recomendación.
- Integración con Plataformas Existentes: Integrar el sistema de recomendación en plataformas académicas existentes, como sistemas de gestión del aprendizaje o bases de datos bibliográficas.
- Pruebas y Despliegue: Realizar pruebas exhaustivas para verificar el funcionamiento correcto del sistema y luego desplegarlo en un entorno de producción para su uso real.



Yan, E. (2021, October 3). *RecSys 2021 - Papers and Talks to Chew on*. eugeneyan.com.
<https://eugeneyan.com/writing/recsys2021/>

Evaluación del Sistema de Recomendación

Las métricas de evaluación son herramientas fundamentales para medir el rendimiento y la calidad de un sistema de recomendación.

- **Precisión:** Mide la proporción de elementos recomendados que son relevantes para el usuario entre todos los elementos recomendados. Se calcula como el número de elementos relevantes recomendados dividido por el número total de elementos recomendados.
- **Recall:** Mide la proporción de elementos relevantes para el usuario que son recomendados correctamente entre todos los elementos relevantes para el usuario. Se calcula como el número de elementos relevantes recomendados dividido por el número total de elementos relevantes en el conjunto de datos.
- **F1-score:** El puntaje F1 es la media armónica de precisión y recall. Proporciona una medida combinada del rendimiento del sistema, teniendo en cuenta tanto la precisión como la exhaustividad. Se calcula como $2 * (\text{precisión} * \text{recall}) / (\text{precisión} + \text{recall})$.
- **RMSE (Root Mean Square Error):** Esta métrica se utiliza comúnmente en sistemas de recomendación basados en valoraciones numéricas. Mide la diferencia entre las valoraciones reales y las valoraciones predichas por el sistema. Se calcula como la raíz cuadrada de la media de los errores al cuadrado.
- **AUC (Area Under the Curve):** Se utiliza principalmente en sistemas de recomendación de clasificación binaria, donde el sistema clasifica los elementos como relevantes o no relevantes. Mide la capacidad del sistema para clasificar correctamente los elementos. Cuanto mayor sea el valor de AUC, mejor será el rendimiento del sistema.
- **MAP (Mean Average Precision):** Esta métrica calcula el promedio de la precisión en los diferentes puntos de recuperación. Es particularmente útil cuando se evalúan sistemas de recomendación que proporcionan una lista clasificada de elementos recomendados.
- **NDCG (Normalized Discounted Cumulative Gain):** Es una métrica que mide la calidad de un conjunto de recomendaciones clasificadas. Toma en cuenta tanto la relevancia de los elementos recomendados como su posición en la lista de recomendaciones.

Consideraciones lingüísticas

En un sistema de recomendación académica, las consideraciones lingüísticas son fundamentales para garantizar una experiencia satisfactoria para los usuarios.

- Es importante tener en cuenta el idioma preferido de los usuarios al presentar recomendaciones y contenido. Por ejemplo, si un usuario prefiere recibir recomendaciones en español, el sistema debe priorizar el contenido en ese idioma.
- Si la audiencia del sistema de recomendación es multilingüe, el sistema debe ser capaz de manejar múltiples idiomas y ofrecer recomendaciones en el idioma preferido de cada usuario. Esto puede requerir la implementación de modelos de procesamiento de lenguaje natural capaces de trabajar con diferentes idiomas.
- Para sistemas que operan en múltiples idiomas, la capacidad de traducción automática puede ser útil para ofrecer recomendaciones en el idioma preferido del usuario, incluso si el contenido original está en otro idioma.
- Además del idioma, el sistema debe considerar las diferencias culturales al recomendar contenido. Lo que puede ser relevante en un contexto cultural puede no serlo en otro. Por lo tanto, el sistema debe adaptar las recomendaciones para que sean culturalmente relevantes para cada usuario.
- Es útil etiquetar el contenido con el idioma en el que está disponible para que el sistema pueda tomar decisiones informadas al recomendar contenido a los usuarios. Esto puede ayudar a evitar recomendaciones inapropiadas en idiomas que el usuario no comprende.
- El sistema debe permitir a los usuarios personalizar su idioma preferido y ajustar las configuraciones lingüísticas según sus necesidades. Esto garantiza una experiencia personalizada y cómoda para cada usuario.

Consideraciones Éticas

- Es fundamental respetar la privacidad de los usuarios y proteger sus datos personales. El sistema debe cumplir con las leyes y regulaciones de protección de datos, y garantizar que la información del usuario se maneje de manera segura y confidencial.
- El sistema debe ser transparente en cuanto a cómo se recopilan, almacenan y utilizan los datos de los usuarios. Los usuarios deben ser informados claramente sobre cómo se generan las recomendaciones y qué información se utiliza para ello.
- Es importante garantizar que el sistema de recomendación sea neutral y no esté sesgado hacia ciertos usuarios, grupos o tipos de contenido. Se deben implementar medidas para evitar la amplificación de sesgos existentes en los datos o algoritmos.
- El sistema debe promover la diversidad y la inclusión al recomendar contenido académico. Se deben tomar medidas para garantizar que se recomienden una variedad de perspectivas, autores y temas, y que se evite la exclusión de ciertos grupos o disciplinas.
- Es importante realizar una evaluación periódica del sistema para identificar posibles sesgos, problemas éticos o impactos negativos. Se deben establecer mecanismos de supervisión y retroalimentación para abordar cualquier problema que surja.

Limitante de los sistemas de recomendación

- Los sistemas de recomendación tienden a recomendar contenido similar al que un usuario ha interactuado en el pasado, lo que puede crear un sesgo y una "burbuja de filtro", donde los usuarios solo ven contenido que ya conocen y limita su exposición a nuevas ideas y perspectivas.
- Si los datos de interacción usuario-ítem son escasos o incompletos, los sistemas de recomendación pueden tener dificultades para generar recomendaciones precisas y relevantes, especialmente en contextos donde hay una amplia variedad de elementos disponibles.
- Los sistemas de recomendación pueden reflejar sesgos demográficos y culturales presentes en los datos de entrenamiento, lo que puede llevar a recomendaciones sesgadas que no representan adecuadamente la diversidad de los usuarios.
- Algunos modelos de recomendación, especialmente los basados en aprendizaje automático, pueden ser difíciles de interpretar, lo que dificulta la comprensión de cómo se generan las recomendaciones y por qué se presentan ciertos elementos a los usuarios.
- Los sistemas de recomendación tienden a priorizar la precisión y la relevancia en las recomendaciones, lo que puede limitar la capacidad de descubrimiento de contenido nuevo e inesperado, conocido como la falta de serendipia.

Conclusión

- Los sistemas de recomendación han demostrado ser herramientas poderosas con una amplia gama de aplicaciones prácticas en diversos campos.
- Estos sistemas facilitan la exploración y el descubrimiento de contenido relevante para los usuarios al mejorar la eficiencia y la personalización de la experiencia de usuario.
- Estos sistemas se han convertido en una parte integral de la infraestructura digital moderna, permitiendo a las empresas ofrecer recomendaciones personalizadas a sus clientes, mejorar la retención de usuarios y aumentar las ventas.
- A medida que la tecnología continúa avanzando y la cantidad de datos disponibles aumenta, se espera que los sistemas de recomendación desempeñen un papel aún más crucial en la forma en que interactuamos con la información y tomamos decisiones en el mundo digital.

Referencias