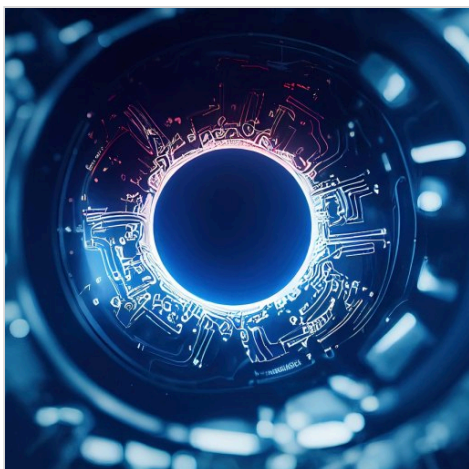


L U M B R E



Enfriamiento inicial y la compensación entre explorar y explotar

Grado: 7

Materia: Sistemas de Recomendación · Inteligencia Artificial

11/10/2026

Antes de empezar

Qué recomendar cuando no hay historia y cuánto cuesta aprender: contenido como andamio, popularidad segmentada, bandidos epsilon-greedy, UCB y Thompson, e interleaving.

QUÉ APRENDERÁS

- Diseñar la política de arranque de un recomendador: puentes para ítem y usuario fríos, bandidos para explorar y evaluación con propensión.

Enfriamiento inicial y la compensación entre explorar y explotar

VIDEO

Cold start en sistemas de recomendación

Video explicativo (inglés): el problema del arranque en frío en recomendadores y las estrategias para resolverlo sin historial de usuario.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=vA96p5zFbSY>

Aprender con usuarios de por medio

Un motor de recomendación llega a un catálogo donde nadie tiene historial y a un usuario nuevo sin gusto conocido. La frialdad inicial y el costo de aprender son el problema real del área; la precisión en datos maduros es la parte fácil.

El **enfriamiento inicial** es la imposibilidad de recomendar bien donde no hay datos: ítem nuevo, usuario nuevo o sistema nuevo. La solución combina puentes (contenido, popularidad, conocimiento del dominio) con **exploración**: mostrar opciones subóptimas a propósito para generar la información que falta. Entre explorar y explotar media una compensación que se gestiona con bandidos.

Las tres frialdades

- Del ítem: productos recién publicados sin una sola interacción.
- Del usuario: cuentas nuevas sin historial que factorizar.
- Del sistema: un catálogo o región donde el modelo entrenado no tiene datos.

Puentes para el ítem frío

Sin interacciones, el contenido habla: metadatos, texto, incrustaciones de portada o descripción, y similitud con ítems ya conocidos. Un **modelo de contenido** que predice afinidad a partir de atributos (no de cofrecuencia) sirve de andamio hasta que las señales colaborativas maduran; el puente más honesto sigue siendo popularidad segmentada por contexto, con su sesgo de exposición a cuestas.



La frialdad se cruza con puentes mientras la exploración acumula evidencia para el colaborativo.

Puentes para el usuario frío

- Onboarding: preguntar gustos explícitos al registrar, con costo de fricción.
- Demografía y dispositivo como rasgos iniciales del contexto.
- Transferencia desde otra región del producto con historia similar.

Bandidos: la contabilidad de explorar

El **bandido multi-brazo** formaliza la decisión: cada recomendación es jalar una palanca cuyo pago es incierto pero aprendible. **Epsilon-greedy** explota y, con probabilidad epsilon, prueba al azar; **UCB** elige optimistamente según el margen superior de su incertidumbre; **Thompson** muestrea de la distribución posterior de cada brazo. Con contexto (LinUCB, bosques condicionales sobre rasgos), la exploración se dirige a los usuarios donde más vale.

El costo de aprender

Explorar gasta la paciencia del usuario: cada impresión malgastada es clics que no llegaron y, en plataformas de suscripción, cancelaciones. La tasa de exploración se fija contra el **regreto** esperado —lo que se deja de ganar por no saber— y se retira cuando la incertidumbre posterior ya no justifica el costo.

Evaluar lo que no se mostró

- Offline: ranking en logs con métricas de orden, con sesgo de posición y de selección.
- Inverse propensity scoring: reponderar los clics observados por la probabilidad con que se mostraron.
- Online: A/B real o interleaving (mezclar resultados de dos motores en una sola lista) para detectar diferencias finas con poca muestra.

Cada frialdad pide su puente; ninguno elimina la necesidad de aprender.

Situación	Puente razonable
Ítem sin interacciones	similaridad por contenido o embedding
Usuario sin historial	onboarding + demografía + popularidad local
Mercado nuevo	transferencia desde mercado vecino
Cola larga con duda	exploración dirigida por bandido contextual

Errores frecuentes

- Esperar que el filtrado colaborativo resuelva el ítem nuevo: no tiene con qué.
- Explorar sin medir el costo: un epsilon fijo en usuarios de pago alto.
- A/B con dos motores y un mes: la varianza del efecto se traga diferencias reales.
- Evaluar con el mismo log que generó la recomendación: confirmas tu sesgo.

Ejemplo resuelto: plataforma con 200 títulos nuevos al día

1. Embedding de metadatos y portada para posicionar el título en espacio de contenido.
2. Recomendador híbrido: score colaborativo cuando hay historia, de contenido cuando no.
3. Thompson sampling por segmento: explorar donde la posterior del nuevo título es ancha.
4. Registrar propensión de exposición para poder evaluar offline sin sesgo.
5. Cortar la exploración cuando la cola posterior del título se estrecha.

¿Para qué sirve en la realidad?

News feeds, e-commerce y catálogos de streaming dedican más ingeniería a la frialdad inicial y la exploración que al algoritmo de moda: los datos fríos son la frontera donde se decide si un usuario nuevo regresa.

1. ¿Qué caracteriza al problema del enfriamiento inicial del ítem?

- ☐ El ítem es popular de más
- ☐ No hay interacciones que permitan al colaborativo estimar afinidad
- ☐ El modelo es demasiado grande
- ☐ Los usuarios no hacen clic

2. El muestreo de Thompson decide qué brazo explorar usando el margen superior de una estimación puntual optimista.

- ☐ Verdadero
- ☐ Falso

3. La técnica de evaluación offline que repondera cada clic por la inversa de la probabilidad con que se mostró el ítem se llama inverse _____ scoring.

4. Clasifica cada estrategia según el tipo de frialdad que atiende.

Elementos:

- similaridad por embedding de metadatos
- preguntas de onboarding
- popularidad segmentada por contexto
- transferencia desde mercado vecino
- predicción de afinidad por atributos

Ítem frío	. . .
Usuario frío	. . .
Ambos	. . .

5. Une cada bandido con su regla de decisión.

5.	Epsilon-greedy	-----
5.	UCB	-----
5.	Thompson	-----
5.	Regreto	-----

Opciones: explotar y con probabilidad fija probar al azar · optimismo ante la incertidumbre de la cota · muestrear la posterior de cada brazo · lo que se paga por no saber

ACTIVIDAD INTERACTIVA DISPONIBLE EN
LUMBRE

6. Enfriamiento y exploración



Enfriamiento y bandidos

Frialdad y exploración

Frío

ítem nuevo

usuario nuevo

mercado nuevo

Puentes

contenido

popularidad

onboarding

Bandidos

epsilon greedy

UCB

Thompson

Evaluación

IPS

interleaving

El mapa mental interactivo es interactivo en la versión online.



Bandido multi-brazo (Wikipedia) https://en.wikipedia.org/wiki/Multi-armed_bandit

Formalización del regret y las reglas epsilon-greedy, UCB y Thompson.

Enfriamiento inicial en recomendación (Wikipedia) [https://en.wikipedia.org/wiki/Cold_start_\(recommender_systems\)](https://en.wikipedia.org/wiki/Cold_start_(recommender_systems)) Tipos de frialdad y estrategias de ataque estándar del área.

7. Una app de recetas lanza 50 platos nuevos por semana y sus usuarios son mayormente recurrentes. Diseña la política de arranque: qué puente usa cada frialdad, qué algoritmo de exploración con qué presupuesto, y cómo sabrías cuándo retirar la exploración.

Referencia curricular

Grado:	7
Materia:	Sistemas de Recomendación · Inteligencia Artificial

Versión interactiva completa (con actividades, videos y autoevaluación):



<https://univ.lumbre.lat/object/cdia-s7-recs-005-cold-start-y-exploracion>

Material educativo generado desde Lumbre · lumbre.lat