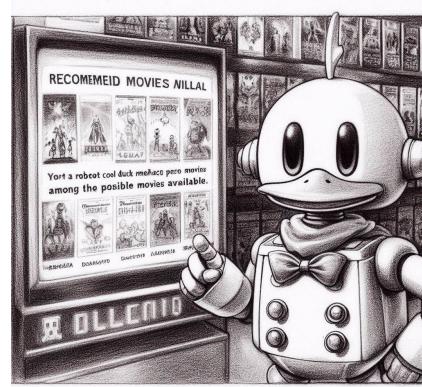


Introduction to recommender systems



AI frameworks

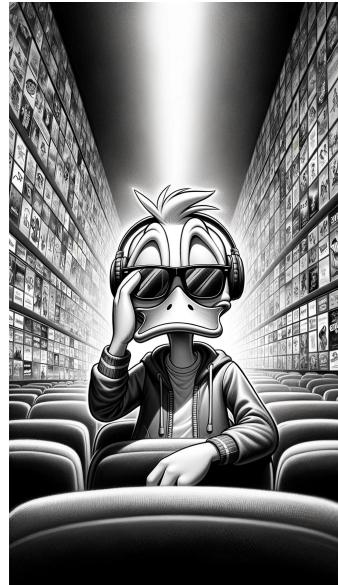
Course: 6 videos

- Introduction
- Popularity and content based methods
- Collaborative filtering
- Matrix factorization
- Neural collaborative filtering
- Evaluating recommender systems

Introduction

Objectives

- Help users to match with the best items
- Ease information overload



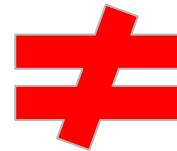
Objectives:

- Information filtering able to predict appetency.



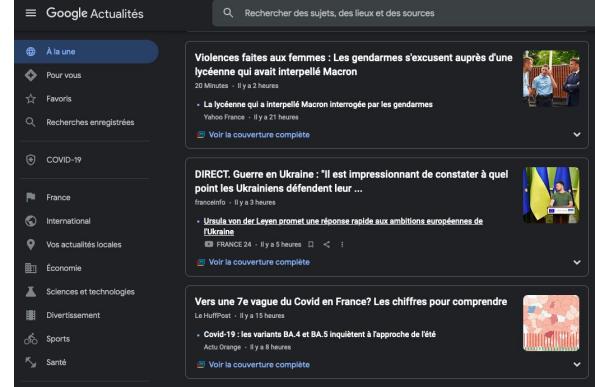
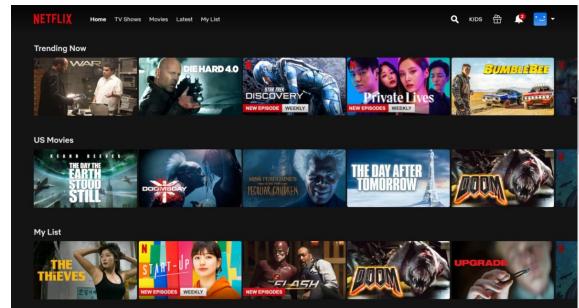
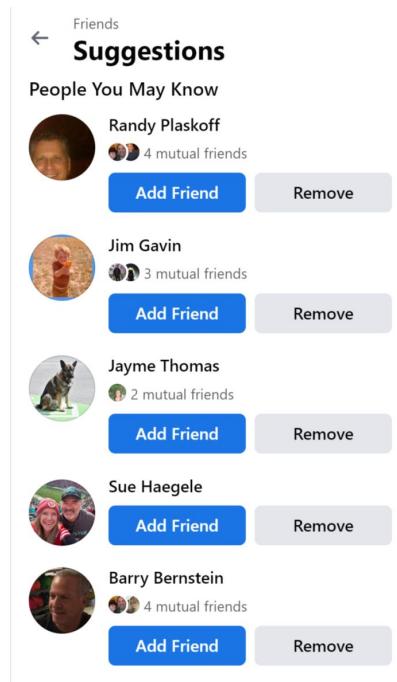
Objectives:

- Information filtering able to predict appetency.
- Different of search



Applications:

- E-commerce
 - Media streaming services
 - News and content aggregation websites
 - Social networks



Applications:

- Netflix => 2/3 of the movies watched are recommended
- Google => news recommendations improved click-through rate (CTR) by 38%
- Amazon => 35% of sales come from recommendations

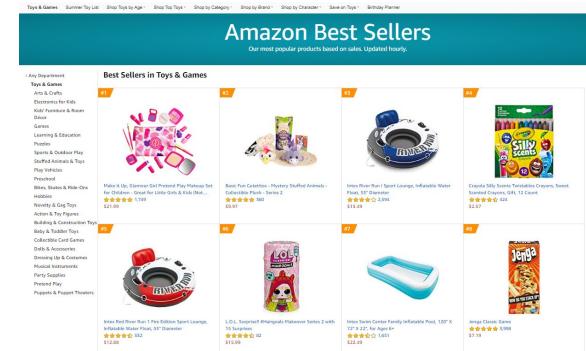
History and evolution

- Rule based systems
- Xerox Palo Alto Research => Tapestry 1992 collaborative filtering Mails
- GroupLens system => 1992 Matrix filling problem News
- Netflix prize 2006-2009
- Lots of side information (item-related data, context, social-networks,...)
- Sequence-aware recommendations

Types of recommender systems

Types of recommender systems

- Popularity



Types of recommender systems

- Popularity
- Content based



Hopefully you love the Unearth Women brand as much as we have loved creating it. Our mission has always been to support women, lift their voices and, simply put, *unearth* women's stories. This female-designed ceramic coffee mug features the Unearth Women logo printed across the mug. The mug measures at a height of 3.85" (9.8 cm) and diameter of 3.35" (8.5 cm) and is microwave and dishwasher safe.

As always, every purchase from the Unearth Women store helps support our mission to pay our female content creators fairly and keep our platform growing.

The screenshot shows the 'Best Sellers in Toys & Games' section of the Amazon website. The products listed are:

- Hula Hoop - Children's Glitter Cat Printed Play Hula Hoop Set for Children - Great for Little Girls & Kids (Ages 3+)
- Basic Fun Cuddlykins - Mystery Stuffed Animals Collection Plush - Series 2
- Inflatable Water Pool - 10' x 2' x 2' for Ages 6+
- Candy Milk - Frootie Tootsie Candy, Sweet Nummy Caramel, GRL, 12oz
- L.O.L. Surprise! Whirligigz Plushie Series 2 with 10 Surprises!
- LOL Surprise! Holographic Fashion Doll with 15 Surprises!
- Inflatable Family Swimming Pool, 130" X 72" X 22" for Ages 6+
- Jenga Classic Game

Types of recommender systems

- Popularity
- Content based
- Collaborative filtering



Hopefully you love the Unearth Women brand as much as we have loved creating it. Our mission has always been to support women, lift their voices and, simply put, *unearth* women's stories. This female-designed ceramic coffee mug features the Unearth Women logo printed across the mug. The mug measures at a height of 3.85" (9.8 cm) and diameter of 3.35" (8.5 cm) and is microwave and dishwasher safe.

As always, every purchase from the Unearth Women store helps support our mission to pay our female content creators fairly and keep our platform growing.

Amazon Best Sellers
Our most popular products based on sales. Updated hourly.

Best Sellers in Toys & Games

Rank	Product	Rating	Price
1	Flamee Cat Pretend Play Pet Set	4.5	\$19.99
2	Mystery Stuffed Animals Collection Plush - Series 2	4.5	\$19.99
3	Inflatable Water Pool, 17' x 12' x 4'	4.5	\$11.99
4	Entertainment Weekly Superhero Trading Cards, Sweet	4.5	\$12.99
5	Jenga Classic Game	4.5	\$19.99



Types of recommender systems

- Popularity
- Content based
- Collaborative filtering
- Graph-based



Hopefully you love the Unearth Women brand as much as we have loved creating it. Our mission has always been to support women, lift their voices and, simply put, *unearth* women's stories. This female-designed ceramic coffee mug



Types of recommender systems

- Popularity
- Content based
- Collaborative filtering
- Graph-based
- Hybrid



Hopefully you love the Unearth Women brand as much as we have loved creating it. Our mission has always been to support women, lift their voices and, simply put, *unearth* women's stories. This female-designed ceramic coffee mug



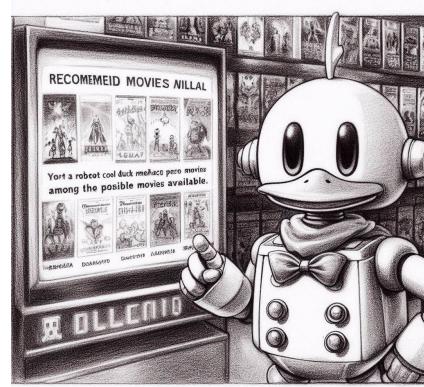
Amazon Best Sellers
Our most popular products based on sales. Updated hourly.

Best Sellers in Toys & Games

Rank	Product	Price	Rating
1	Hatch N' Go! Glamour Cat Preemie Playset	\$24.99	4.5
2	Basic Fun Cutiekins - Mystery Stuffed Animals Collection Pack - Series 2	\$14.99	4.5
3	Inflatable Water Park - 10x10ft	\$11.99	4.5
4	Entertainment Weekly Superhero Crayons, Sweet Superhero Crayons, GRL, 12 Count	\$2.17	4.5
5	Jenga Classic Game	\$1.99	4.5



Introduction to recommender systems



AI frameworks

Types of recommender systems

- Popularity
- Content based
- Collaborative filtering
- Graph-based
- Hybrid



Amazon Best Sellers

Our most popular products based on sales. Updated hourly.

Best Sellers in Toys & Games

Rank	Product	Rating	Price
1	HABA Haba Baby - Play Pretend Play Pretend Set	4.5	\$29.99
2	Basic Fun Cutiekins - Mystery Stuffed Animals Collection Pack - Series 2	4.5	\$14.99
3	Inflatable Water Pool, 117" x 72" x 24"	4.5	\$11.49
4	Entertainment Earth Exclusive Jenga Game	4.5	\$22.99
5	Kids River Run 1-Person Inflatable Water	4.5	\$12.99
6	L.O.L. Surprise! Whirligig Maternity Series 2 with 10 Surprises!	4.5	\$15.99
7	Inflatable Pool Lounger, Inflatable Water Pool, 117" x 72" x 24" for Ages 6+	4.5	\$22.49
8	Jenga Classic Game	4.5	\$19.99



Popularity based recommendations

Popularity based:

- Rank items and recommend most popular



Amazon Best Sellers

Our most popular products based on sales. Updated hourly.

Best Sellers in Toys & Games

#	Product	Description	Rating	Price
#1	Make It Up, Glamour Girl Pretend Play Makeup Set for Children - Great for Little Girls & Kids (Not ...	\$21.99 4.5 stars 1,740 reviews	4.5	\$21.99
#2	Basic Fun Cutetitos - Mystery Stuffed Animals - Cinnabon Plush - Series 2	\$9.97 4.5 stars 360 reviews	4.5	\$9.97
#3	Intex River Run I Sport Lounge, Inflatable Water Float, 31" Diameter	\$13.49 4.5 stars 2,294 reviews	4.5	\$13.49
#4	Crayola Silly Scents Twistables Crayons, Sweet Scented Crayons, Gift, 12 Count	\$2.57 4.5 stars 424 reviews	4.5	\$2.57
#5	Fun 1 Fire Edition Sport Lounge, Inflatable Pool Float, 53" Diameter	\$15.99 4.5 stars 2 reviews	4.5	\$15.99
#6	L.O.L. Surprise! #Hargals Makeover Series 2 with 7 Surprises	\$15.99 4.5 stars 82 reviews	4.5	\$15.99
#7	Intex Swim Center Family Inflatable Pool, 120" X 72" X 22", for Ages 6+	\$22.49 4.5 stars 1,051 reviews	4.5	\$22.49
#8	Jenga Classic Game	\$7.19 4.5 stars 5,958 reviews	4.5	\$7.19



Popularity based:

- Rank items and recommend most popular
- Can be specified by category or time period



Any Department
Toys & Games
Arts & Crafts
Electronics for Kids
Kid's Furniture & Room Decor
Games
Learning & Education
Puzzles
Sports & Outdoor Play
Stuffed Animals & Toys
Toy Vehicles
Preschool
Bikes, Skates & Ride-Ons
Hobbies
Novelty & Gag Toys
Puzzles & Trivia Games
Building & Construction Toys
Baby & Toddler Toys
Collectible Card Games
Dolls & Accessories
Dressing Up & Costumes
Musical Instruments
Party Supplies
Princess Play
Puppets & Puppet Theaters

Best Sellers in Toys & Games

Rank	Product Image	Product Name	Rating	Reviews	Price
#1		Make It Up, Glamour Girl Pretend Play Makeup Set for Children - Great for Little Girls & Kids (Not Real)	4.5	1,740	\$21.99
#2		Basic Fun Cutetitos - Mystery Stuffed Animals - Colorful Plush - Series 2	4.5	360	\$9.97
#3		Intex River Run I Sport Lounge, Inflatable Water Float, 31" Diameter	4.5	2,294	\$13.49
#4		Crayola Silly Scents Twistables Crayons, Sweet Scented Crayons, Gift, 12 Count	4.5	424	\$7.37
#5		Fun 1 Fire Edition Sport Lounge, Float, 53" Diameter	4.5	2	\$15.99
#6		L.O.L. Surprise! #Hargals Makeover Series 2 with 7 Surprises	4.5	82	\$15.99
#7		Intex Swim Center Family Inflatable Pool, 120" X 72" X 22", for Ages 6+	4.5	1,051	\$22.49
#8		Jenga Classic Game	4.5	5,998	\$7.19



Popularity based:

- Rank items and recommend most popular
- Can be specified by category or time period

Advantages:

- Easy to set up
- No information on users needed
- Good cold start
- Good baseline



Any Department
Toys & Games
Arts & Crafts
Electronics for Kids
Kid's Furniture & Room Decor
Games
Learning & Education
Puzzles
Sports & Outdoor Play
Stuffed Animals & Toys
Toy Vehicles
Preschool
Bikes, Skates & Ride-Ons
Hobbies
Novelty & Gag Toys
Puzzles & Trivia Games
Building & Construction Toys
Baby & Toddler Toys
Collectible Card Games
Dolls & Accessories
Dressing Up & Costumes
Musical Instruments
Party Supplies
Princess Play
Puppets & Puppet Theaters

Best Sellers in Toys & Games

Rank	Product	Description	Rating	Reviews	Price
#1	Make It Up, Glamour Girl Pretend Play Makeup Set for Children - Great for Little Girls & Kids (Not ...)	2.5 stars, 1,740 reviews	\$21.99		
#2	Basic Fun Cutetitos - Mystery Stuffed Animals - Cuteness Plush - Series 2	3.5 stars, 360 reviews	\$9.97		
#3	Intex River Run I Sport Lounge, Inflatable Water Float, 31" Diameter	3.5 stars, 2,294 reviews	\$13.49		
#4	Crayola Silly Scents Twistables Crayons, Sweet Scented Crayons, Gift, 12 Count	3.5 stars, 424 reviews	\$2.57		
#5	Fun 1 Fire Edition Sport Lounge, Inflatable Water Float, 33" Diameter	2 stars, 2 reviews	\$15.99		
#6	L.O.L. Surprise! #Hargals Makeover Series 2 with 7 Surprises	3 stars, 82 reviews	\$15.99		
#7	Intex Swim Center Family Inflatable Pool, 120" X 72" X 22", for Ages 6+	3 stars, 12 reviews	\$22.49		
#8	Jenga Classic Game	4 stars, 5,958 reviews	\$7.19		

Popularity based:

- Rank items and recommend most popular
- Can be specified by category or time period

Advantages:

- Easy to set up
- No information on users needed
- Good cold start
- Good baseline



Any Department
Toys & Games
Arts & Crafts
Electronics for Kids
Kid's Furniture & Room Decor
Games
Learning & Education
Puzzles
Sports & Outdoor Play
Stuffed Animals & Toys
Toy Vehicles
Preschool
Bikes, Skates & Ride-Ons
Hobbies
Novelty & Gag Toys
Puzzles & Trivia Games
Building & Construction Toys
Baby & Toddler Toys
Collectible Card Games
Dolls & Accessories
Dressing Up & Costumes
Musical Instruments
Party Supplies
Princess Play
Puppets & Puppet Theaters

Best Sellers in Toys & Games

Rank	Product	Description	Rating	Price
#1	Make it Up, Glamour Girl Pretend Play Makeup Set for Children - Great for Little Girls & Kids (Not ...)	\$21.99	4.5 ★★★★☆ 1,740	
#2	Basic Fun Cutetitos - Mystery Stuffed Animals - Cuteness Plush - Series 2	\$9.97	4.5 ★★★★☆ 360	
#3	Intex River Run I Sport Lounge, Inflatable Water Float, 31" Diameter	\$13.49	4.5 ★★★★☆ 2,294	
#4	Crayola Silly Scents Twistables Crayons, Sweet Scented Crayons, Gift, 12 Count	\$7.37	4.5 ★★★★☆ 424	
#5	Run 1 Fire Edition Sport Lounge, Float, 53" Diameter	\$15.99	4.5 ★★★★☆ 82	
#6	L.O.L. Surprise! #Hargals Makeover Series 2 with Fashions	\$12.49	4.5 ★★★★☆ 1,031	
#7	Intex Swim Center Family Inflatable Pool, 120" X 72" X 22", for Ages 6+	\$22.49	4.5 ★★★★☆ 1,031	
#8	Jenga Classic Game	\$7.19	4.5 ★★★★☆ 5,998	

Drawbacks:

- No personalization
- Popular items may end up dominating the recommendations



Content based recommendations

Content based

Goal: Find most similar items based on their characteristics

Content based

Item description:



Marque	Xinlon
Dimensions du produit	7,5L x 6l x 6H centimètres
Tranche d'âge (description)	Enfant
Nombre de pièces	2
Nombre d'articles	2
Nom de la collection	Animal
Nombre d'unités	2 Count
Fabricant	Xinlongyi-Bin



2 x Rubber ducks

Content based

Textual description:



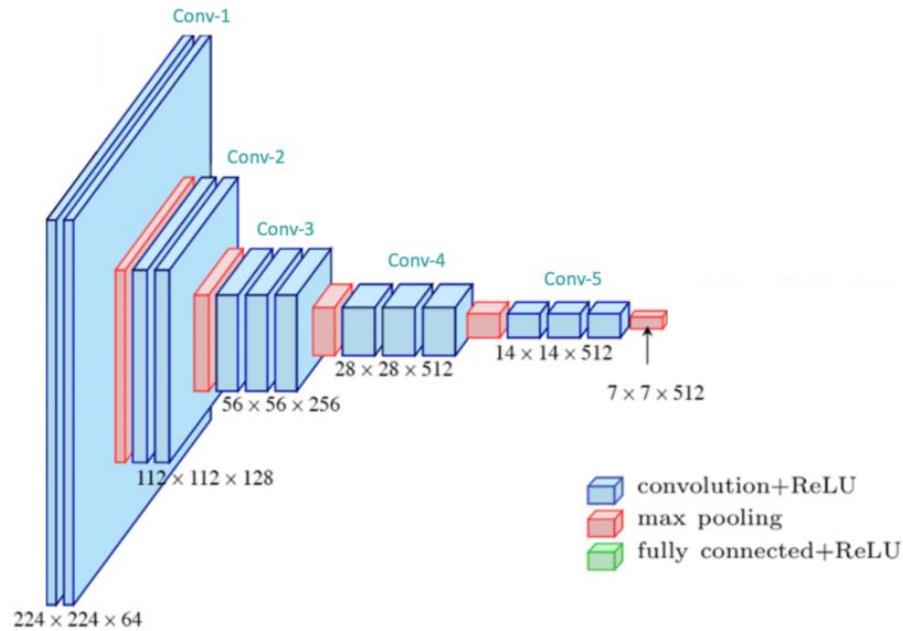
À propos de cet article

- **Matériaux de haute qualité:** les jouets de canard jaune sont fabriqués en silicone sans danger et ne causeront aucun dommage à la peau humaine. Ils ont une surface lisse et un bon toucher, lisses, doux, résistants aux températures élevées et difficiles à décolorer.
- **Beau design:** la décoration de la voiture de canard jaune est équipée de lunettes de soleil, de colliers et de casques sympas. La forme est très intéressante et mignonne, et il est facile d'attirer l'attention des autres. Il convient à la décoration de votre voiture et de votre casque.
- **Assemblage sans bricolage:** Le jouet mignon de canard est livré avec des casques intéressants, des lunettes de soleil, des colliers et du ruban adhésif double face. Vous devez l'assembler vous-même. Vous pouvez utiliser votre imagination pour vous habiller comme vous le souhaitez. Le casque de ce canard est détachable.
- **Contenu de l'emballage:** Le produit comprend 2 canards en caoutchouc, 2 casques de styles différents, 2 hélices rotatives, 2 lunettes de soleil, 2 colliers et 2 autocollants double face super collants.
- **Scène applicable:** ce type de canard de voiture peut également être collé sur des chariots et des vélos pour enfants. Lorsque les enfants voient des canards jaunes sur des poussettes et des vélos et regardent les hélices sur leurs casques tourner, ils se sentiront très heureux et obéissants.

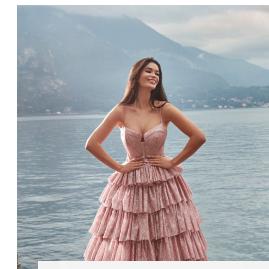
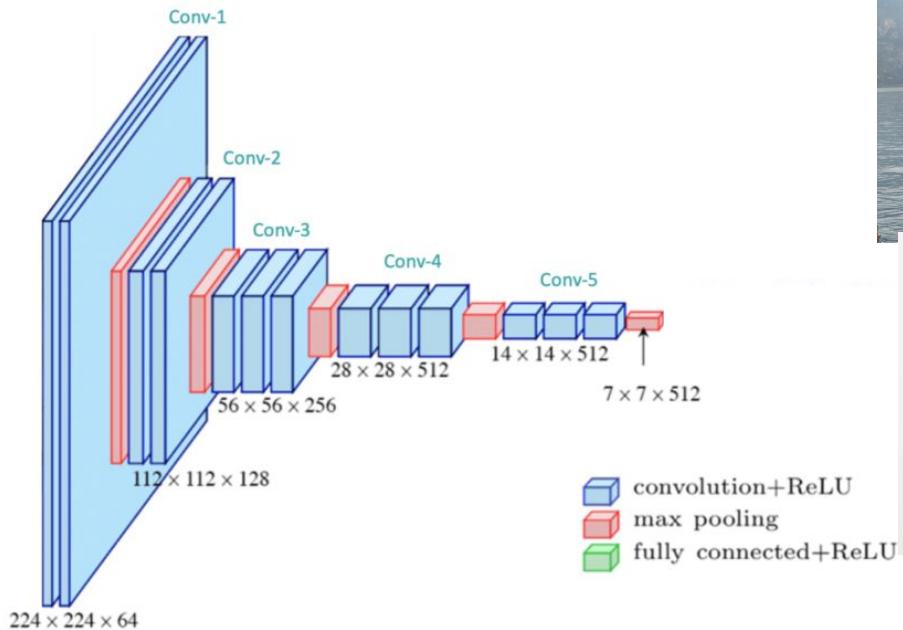


Content based

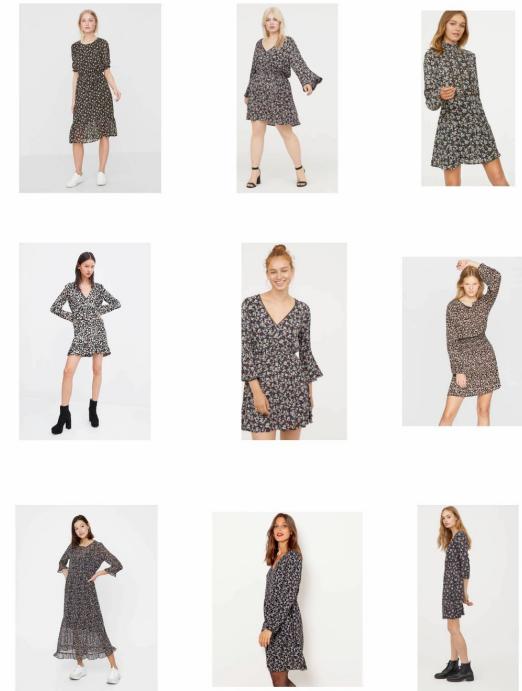
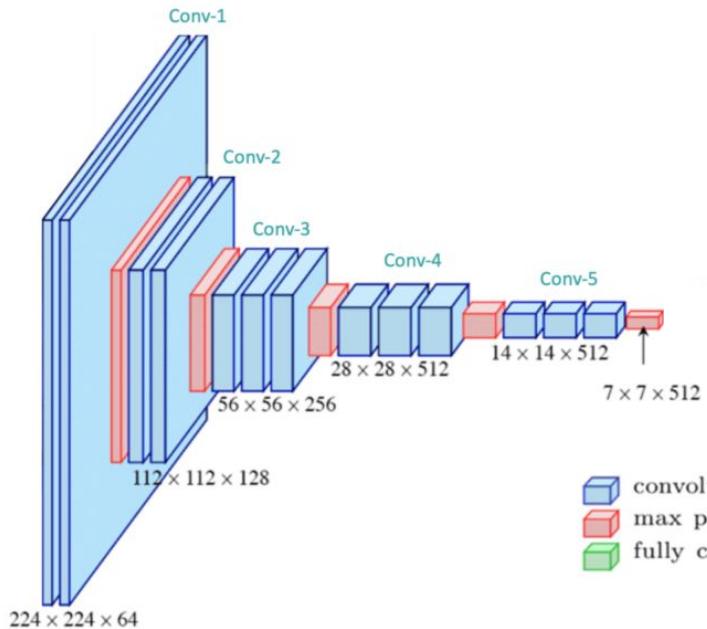
Visual description:



Content based

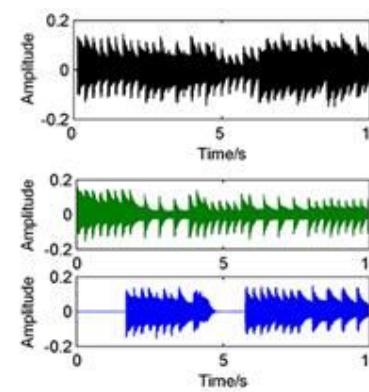
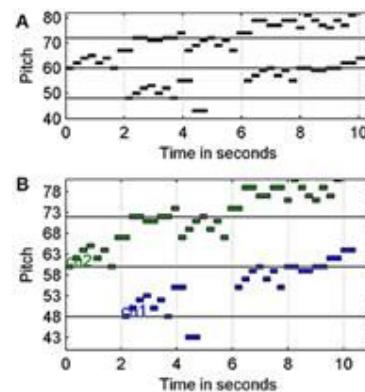


Content based



Content based

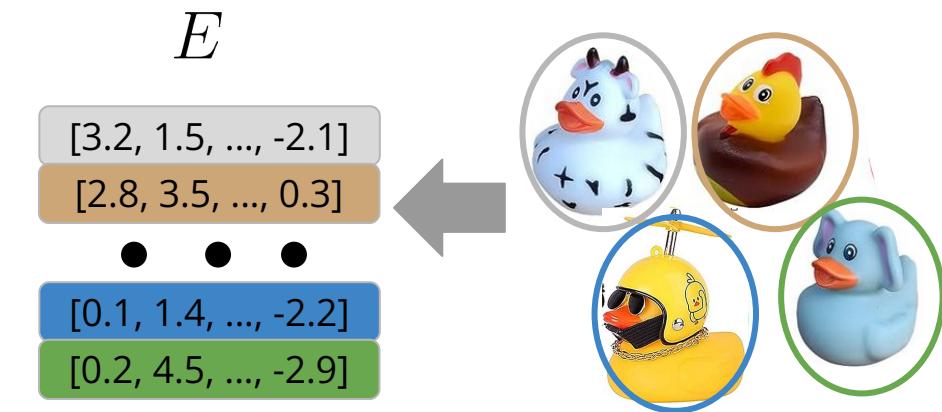
Audio description:



Similarity measures

Similarity measures

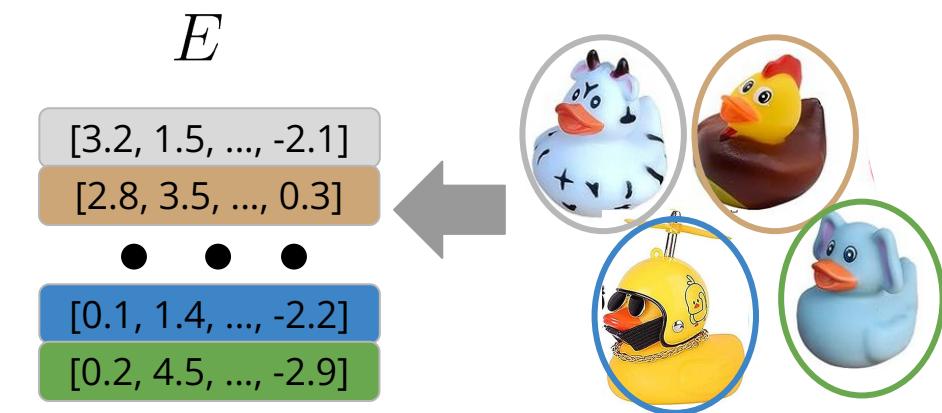
Embedding space E



Similarity measures

Embedding space E

Similarity measure $s : E \times E \rightarrow \mathbb{R}$

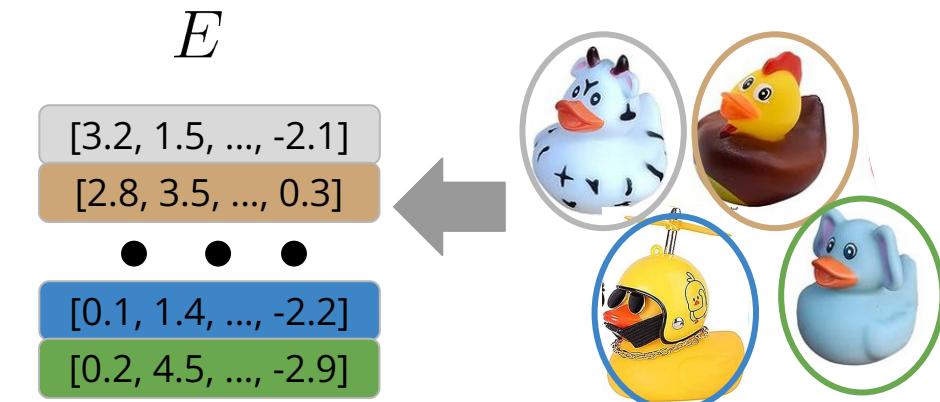
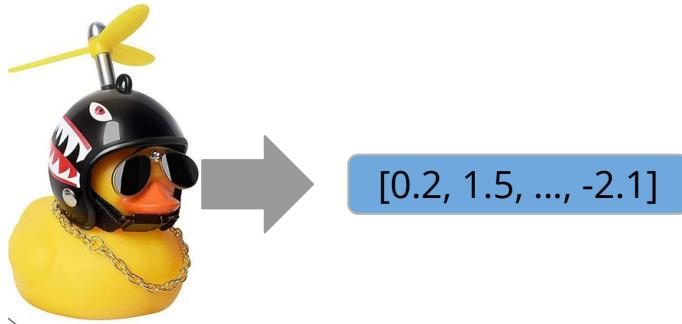


Similarity measures

Embedding space E

Similarity measure $s : E \times E \rightarrow \mathbb{R}$

Given a **query** $q \in E$

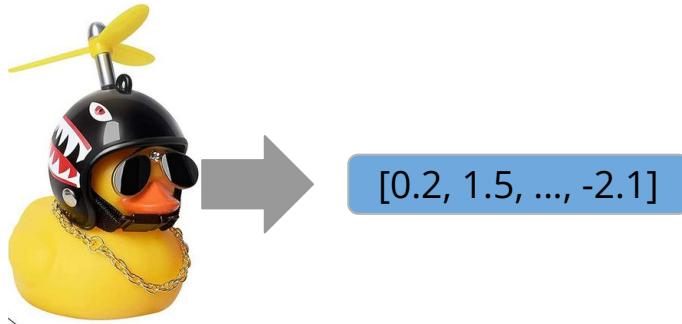


Similarity measures

Embedding space E

Similarity measure $s : E \times E \rightarrow \mathbb{R}$

Given a **query** $q \in E$ the system looks for item embeddings $x \in E$ with **high similarity** $s(q, x)$



E

[3.2, 1.5, ..., -2.1]

[2.8, 3.5, ..., 0.3]

• • •

[0.1, 1.4, ..., -2.2]

[0.2, 4.5, ..., -2.9]



Similarity measures

Embedding space E

Similarity measure $s : E \times E \rightarrow \mathbb{R}$

Given a **query** $q \in E$ the system looks for item embeddings $x \in E$ with **high similarity** $s(q, x)$

- Cosine similarity $s(q, x) = \cos(q, x)$

Similarity measures

Embedding space E

Similarity measure $s : E \times E \rightarrow \mathbb{R}$

Given a **query** $q \in E$ the system looks for item embeddings $x \in E$ with **high similarity** $s(q, x)$

- Cosine similarity $s(q, x) = \cos(q, x)$
- Dot Product $s(q, x) = \langle q, x \rangle = \sum_{i=1}^d q_i x_i = \|x\| \|q\| \cos(q, x)$

Similarity measures

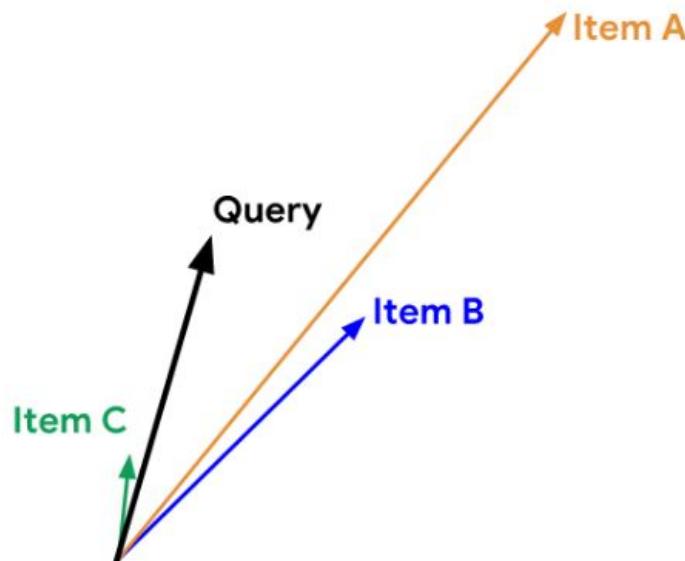
Embedding space E

Similarity measure $s : E \times E \rightarrow \mathbb{R}$

Given a **query** $q \in E$ the system looks for item embeddings $x \in E$ with **high similarity** $s(q, x)$

- Cosine similarity $s(q, x) = \cos(q, x)$
- Dot Product $s(q, x) = \langle q, x \rangle = \sum_{i=1}^d q_i x_i = \|x\| \|q\| \cos(q, x)$
- Euclidian distance $s(q, x) = \|q - x\| = \left[\sum_{i=1}^d (q_i - x_i)^2 \right]^{\frac{1}{2}}$

Similarity



Dot-Product

Query : **Item A > Item B > Item C**

Cosine

Query : **Item C > Item A > Item B**

(-) Euclidean Distance

Query : **Item B > Item C > Item A**

Content based

Advantages:

- No need of user interactions
- Good for cold start
- Easy to scale to large number of users
- Easy to use with new product

Content based

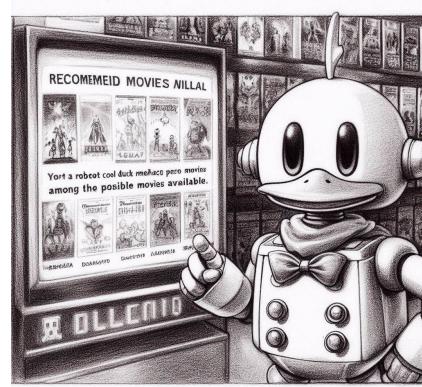
Advantages:

- No need of user interactions
- Good for cold start
- Easy to scale to large number of users
- Easy to use with new product

Drawbacks:

- Hard to extract relevant embeddings
- Embeddings are often handcrafted and require expert knowledge

Introduction to recommender systems



AI frameworks

Types of recommender systems

- Popularity
- Content based
- Collaborative filtering
- Graph-based
- Hybrid



Amazon Best Sellers

Our most popular products based on sales. Updated hourly.

Best Sellers in Toys & Games

Rank	Product	Rating	Price
1	HABA Haba Baby - Play Pretend Play Pretend Set	4.5	\$29.99
2	Basic Fun Cutiekins - Mystery Stuffed Animals Collection Pack - Series 2	4.5	\$14.99
3	Inflatable Water Pool, 117" x 72" x 24"	4.5	\$11.49
4	Crayola Washable Paint, 12 Count	4.5	\$12.19
5	Kids River Run 1-Person Spray Lounge, Inflatable Water Pool, 51" Diameter	4.5	\$12.99
6	L.O.L. Surprise! Whirligig Maternity Series 2 with 10 Surprises	4.5	\$15.99
7	Inflatable Family Sprinkler, 130" x 72" x 24"	4.5	\$22.49
8	Jenga Classic Game	4.5	\$19.99

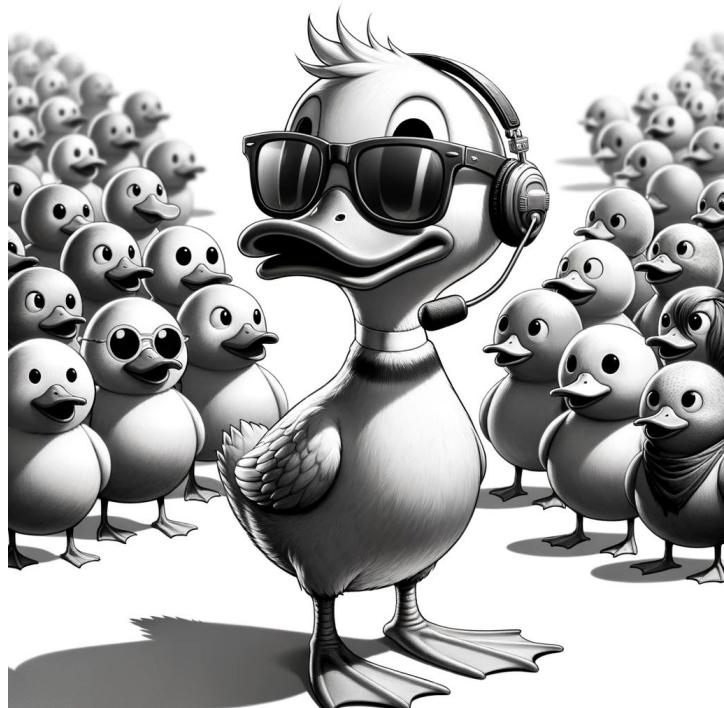


Collaborative filtering

Collaborative filtering

Goal:

Predict the interests of a user using preferences from other users.

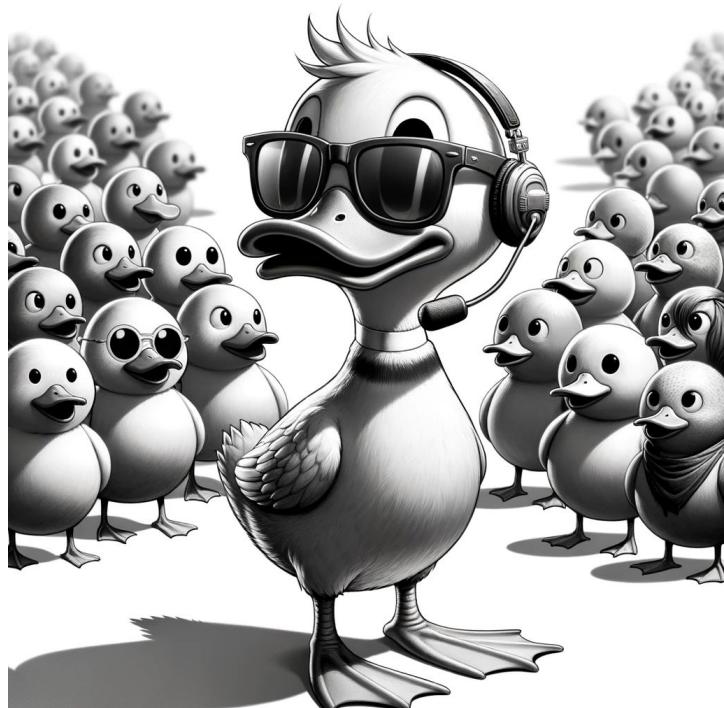


Collaborative filtering

Goal:

Predict the interests of a user using preferences from other users.

- Users x Items interactions
- User based
- Item based
- Matrix factorization



Collaborative filtering

i



Based on interactions
User x Items

Two strategies:

- User-based
- Item-based

<i>u</i>	JURASSIC PARK	The Godfather	The Lord of the Rings	No Time to Die	<i>i</i>
	4	1	4	3	$r_{u,i}$
	1	5	5	4	4
	2	2	2	3	2
	5	5	1	1	1
	4	2	4	3	4
	3	1	4	3	3

User based

Principle:

- Find most similar users
- Estimate rating by the weighted average of similar users

		JURASSIC PARK	The Godfather	Lord of the Rings	NO TIME TO DIE	BATMAN
 <i>u</i>	4	1	4	3	$r_{u,i}$	
	1	5	5	4	4	
	2	2	2	3	2	
	5	5	1	1	1	
	4	2	4	3	4	
	3	1	4	3	3	

User based

Similar users:

- k-nearest neighbors
- Pearson correlation
- Cosine similarity

Neighborhood U



u	Jurassic Park	The Godfather	The Lord of the Rings	Mission Impossible: No Time to Die	Batman Returns	$r_{u,i}$
	4	1	4	3		
	1	5	5	4		4
	2	2	2	3		2
	5	5	1	1		1
	4	2	4	3		4
	3	1	4	3		3

User based

Similar users:

- k-nearest neighbors
- Pearson correlation
- Cosine similarity

Neighborhood U



u	4	1	4	3	$r_{u,i}$
	1	5	5	4	4
	2	2	2	3	2
	5	5	1	1	1
	4	2	4	3	4
	3	1	4	3	3

i

User based

Similar users:

- k-nearest neighbors
- Pearson correlation
- Cosine similarity

Neighborhood U

Estimate rating $r_{u,i}$:

$$r_{u,i} = \frac{1}{N} \sum_{u' \in U} r_{u',i}$$

		Jurassic Park	The Godfather	The Lord of the Rings	Quantum of Solace	Batman Returns
 u	4	1	4	3	$r_{u,i}$	
	1	5	5	4	4	
	2	2	2	3	2	
	5	5	1	1	1	
	4	2	4	3	4	
	3	1	4	3	3	

User based

i



<i>u</i>	4	1	4	3	$r_{u,i}$
	1	5	5	4	4
	2	2	2	3	2
	5	5	1	1	1
	4	2	4	3	4
	3	1	4	3	3

Similar users:

- k-nearest neighbors
- Pearson correlation
- Cosine similarity

Neighborhood U

Similarity measure $\text{sim}(u, u')$

Estimate rating $r_{u,i}$:

$$r_{u,i} = k \sum_{u' \in U} \text{sim}(u, u') r_{u',i}$$

with $k = 1 / \sum_{u' \in U} |\text{sim}(u, u')|$

User based

i



<i>u</i>	4	1	4	3	$r_{u,i}$
	1	5	5	4	4
	2	2	2	3	2
	5	5	1	1	1
	4	2	4	3	4
	3	1	4	3	3

Neighborhood U

Similarity measure $\text{sim}(u, u')$

Estimate rating $r_{u,i}$:

$$r_{u,i} = \overline{r_u} + k \sum_{u' \in U} \text{sim}(u, u') (r_{u',i} - \overline{r_{u'}}) \quad \text{with } \overline{r_u} \text{ the average rating of user } u$$

User based



u	4	1	4	3	$r_{u,i}$
	1	5	5	4	4
	2	2	2	3	2
	5	5	1	1	1
	4	2	4	3	4
	3	1	4	3	3

Drawbacks:

- How to handle new users?
- Does not scale to large real-world scenarios where $|Users| \gg |Items|$

Item based

Principle:

- Use similarity between items
- Recommend items similar with the current item or the user history



u	Jurassic Park	The Godfather	The Lord of the Rings	101 Dalmatians	Batman Returns
	4	1	4	3	$r_{u,i}$
	1	5	5	4	4
	2	2	2	3	2
	5	5	1	1	1
	4	2	4	3	4
	3	1	4	3	3

Item based

Similar items:

- k-nearest neighbors
- Cosine similarity



		JURASSIC PARK	The Godfather	The Lord of the Rings	101 DALMATIANS	BATMAN RETURNS
 u	4	1	4	3		$r_{u,i}$
	1	5	5	4		4
	2	2	2	3		2
	5	5	1	1		1
	4	2	4	3		4
	3	1	4	3		3

Item based

Similar items:

- k-nearest neighbors
- Cosine similarity

Neighborhood I

Similarity measure $\text{sim}(i, i')$

Estimate rating $r_{u,i}$:

$$r_{u,i} = k \sum_{i' \in I} \text{sim}(i, i') r_{u,i'}$$

$$\text{with } k = 1 / \sum_{i' \in I} |\text{sim}(i, i')|$$

u	4	1	4	3	$r_{u,i}$	
	1	5	5	4	4	
	2	2	2	3	2	
	5	5	1	1	1	
	4	2	4	3	4	
	3	1	4	3	3	

Item based

i



Advantages:

- Supposed to be more stable
- Pre-compute pairwise similarities
- Easier to scale

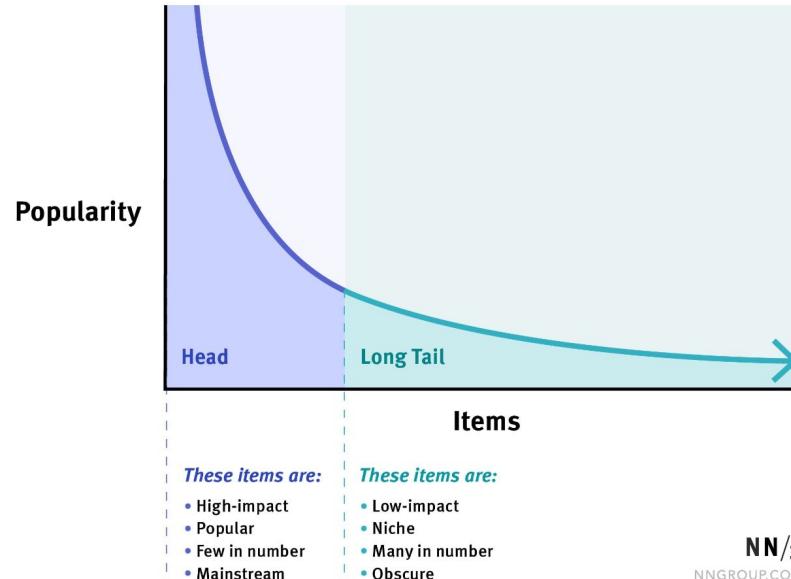
Drawbacks:

- New items
- Items with few interactions

<i>u</i>	4	1	4	3	$r_{u,i}$
	1	5	5	4	2
	2	2	2	3	2
	5	5	1	1	4
	4	2	4	3	?
	3	1	4	3	5

Collaborative filtering

The Long Tail



Explicit vs Implicit Feedbacks



Explicit feedbacks:

- User rate items

	u	4	1	4	3	1
		5	1	5	4	4
		2	2	2	3	2
		5	5	1	1	1
		4	2	4	3	4
		3	1	4	3	3

Explicit vs Implicit Feedbacks



Explicit feedbacks:

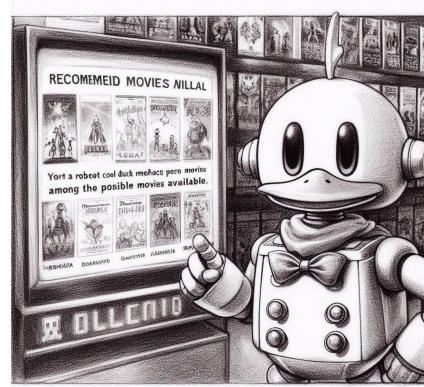
- User rate items

Implicit feedbacks:

- Clicks, watched, bought, ...
- Watch full video, ...
- Stop watching, ...

	<i>u</i>	1	1			1
			1			
				1		
		1		1	1	
		1				1
		1	1			

Introduction to recommender systems



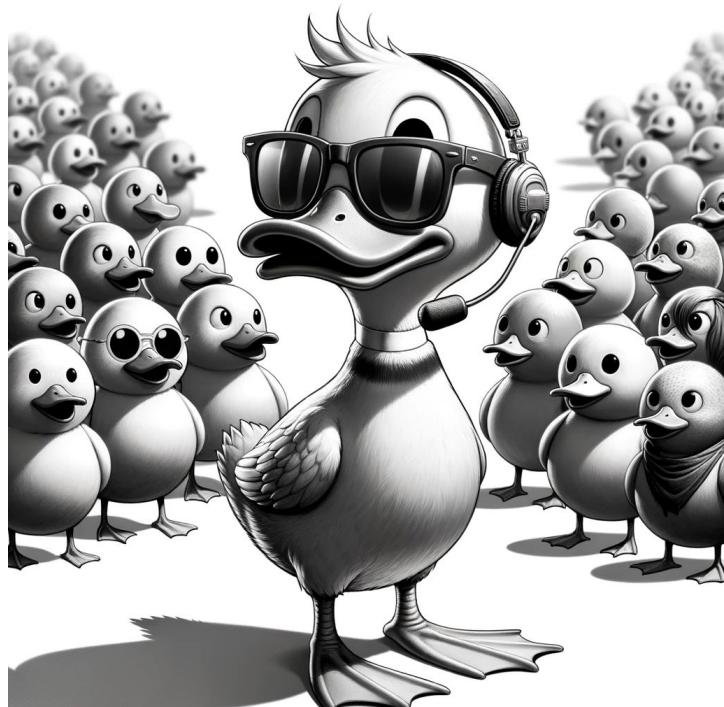
AI frameworks

Collaborative filtering

Goal:

Predict the interests of a user using preferences from other users.

- Users x Items interactions
- User based
- Item based
- Matrix factorization



Matrix factorization

Goal:

Factorize the interaction matrix A



Matrix factorization

Goal:

Factorize the interaction matrix A



Matrix factorization

Goal:

Factorize the interaction matrix A

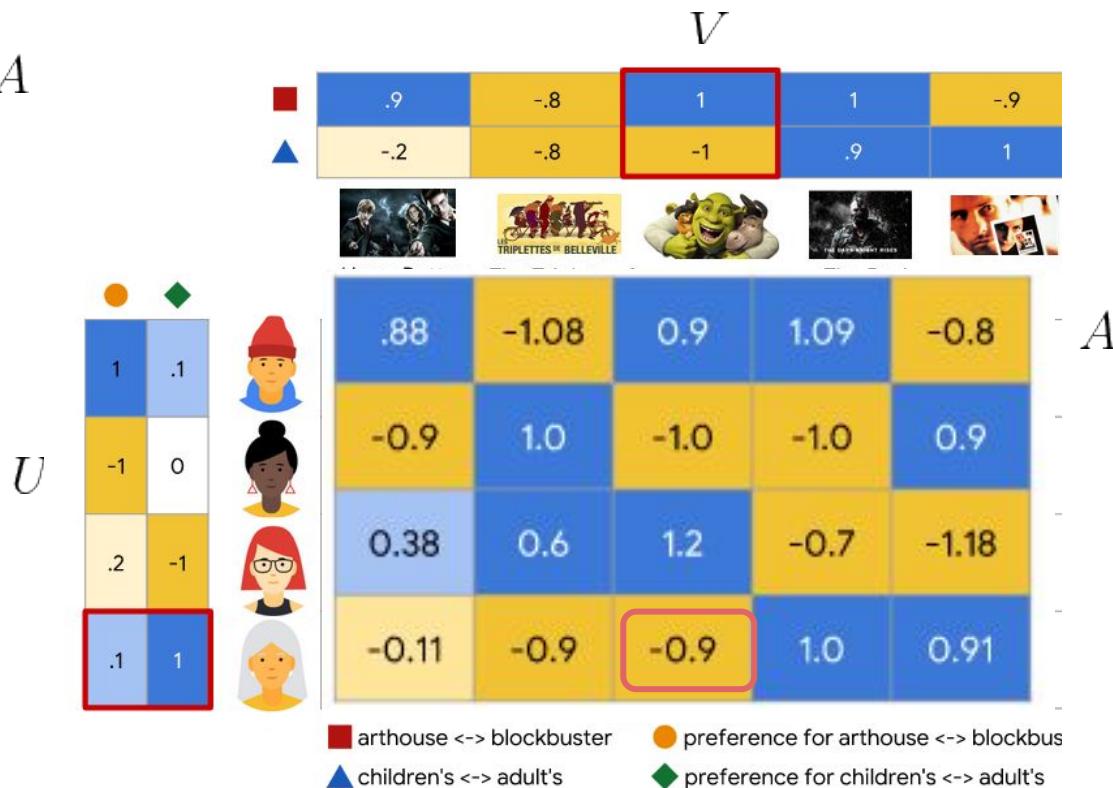


Matrix factorization

Goal:

Factorize the interaction matrix A

- A user embedding matrix U
- An item embedding matrix V

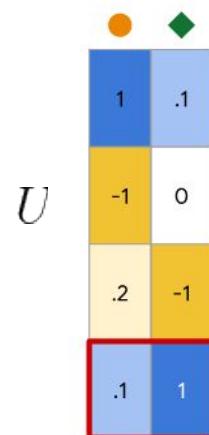


Matrix factorization

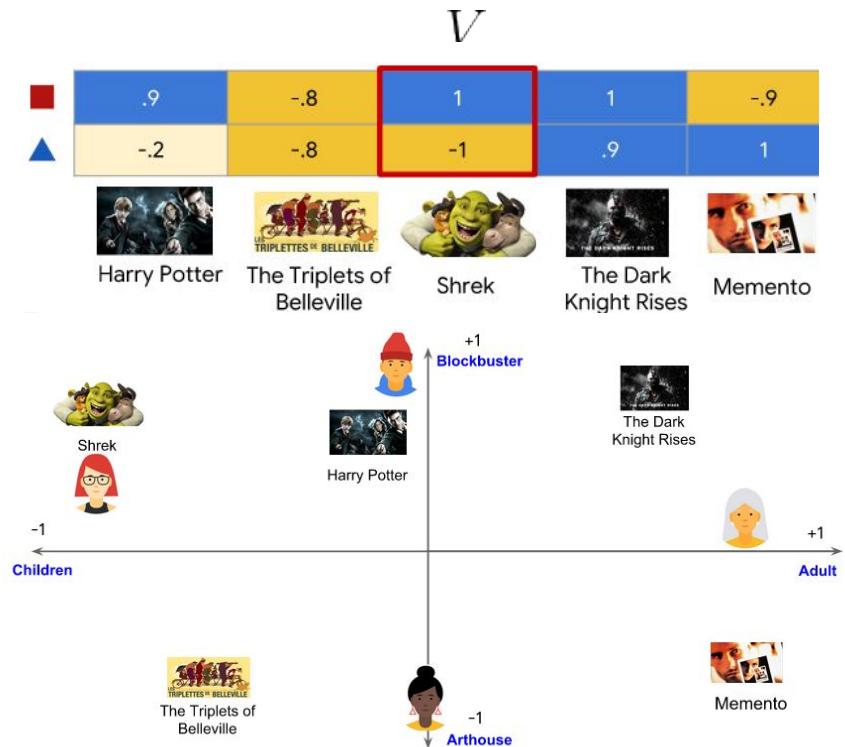
Goal:

Factorize the interaction matrix A

- A user embedding matrix U
- An item embedding matrix V



U



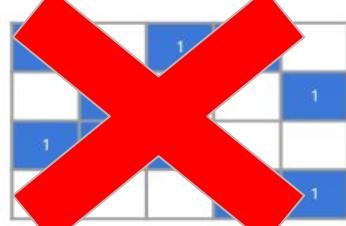
Matrix factorization

Optimization:

- Stochastic gradient descent
- Alternating Least Squares:
 - Fix U and solve for V
 - Fix V and solve for U



Observed Only MF



$$\sum_{(i, j) \in \text{obs}} (A_{ij} - U_i \cdot V_j)^2$$

SVD

1	0	1	1	0
0	1	0	0	1
1	1	1	0	0
0	0	0	1	1

$$= \sum_{(i, j)} |A_{ij} - U_i \cdot V_j|^2$$

Weighted MF

1	0	1	1	0
0	1	0	0	1
1	1	1	0	0
0	0	0	1	1

$$\sum_{(i, j) \in \text{obs}} (A_{ij} - U_i \cdot V_j)^2 + w_0 \sum_{(i, j) \notin \text{obs}} (0 - U_i \cdot V_j)^2$$

Matrix factorization

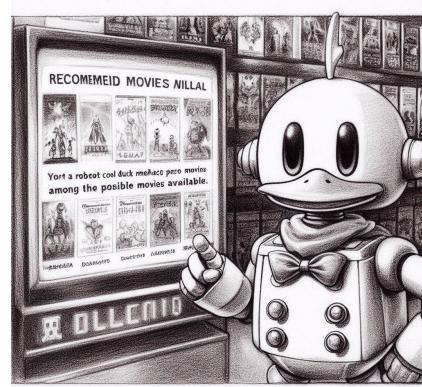
Advantages:

- Can be parallelized (ALS)
- Can be computed offline
- Embeddings can be used for item-item recommendations
- Good for serendipity

Drawbacks:

- Can't handle new items
- Does not include other possible meaningful features

Introduction to recommender systems



AI frameworks

Types of recommender systems

- Popularity
- Content based
- Collaborative filtering
- Graph-based
- Hybrid



Amazon Best Sellers

Our most popular products based on sales. Updated hourly.

Best Sellers in Toys & Games

Rank	Product	Rating	Price
1	HABA Haba Baby - Play Pretend Play Pretend Set	4.5	\$29.99
2	Basic Fun Cutiekins - Mystery Stuffed Animals Collection Pack - Series 2	4.5	\$14.99
3	Inflatable Water Pool, 117" x 72" x 24"	4.5	\$11.49
4	Entertainment Earth Exclusive Jenga Game	4.5	\$22.99
5	Kids River Run 1-Person Inflatable Water	4.5	\$12.99
6	L.O.L. Surprise! Whirligig Maternity Series 2 with 10 Surprises!	4.5	\$12.99
7	Inflatable Pool Lounger, Inflatable Water Pool, 117" x 72" x 24" for Ages 6+	4.5	\$22.49
8	Jenga Classic Game	4.5	\$19.99

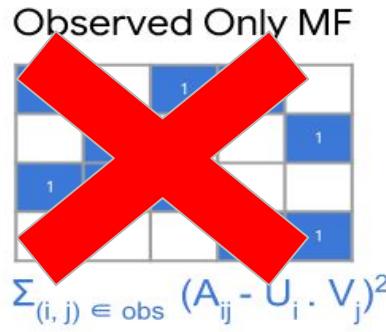


Neural collaborative systems

Matrix factorization

Optimization:

- Stochastic gradient descent
- Alternating Least Squares:
 - Fix U and solve for V
 - Fix V and solve for U



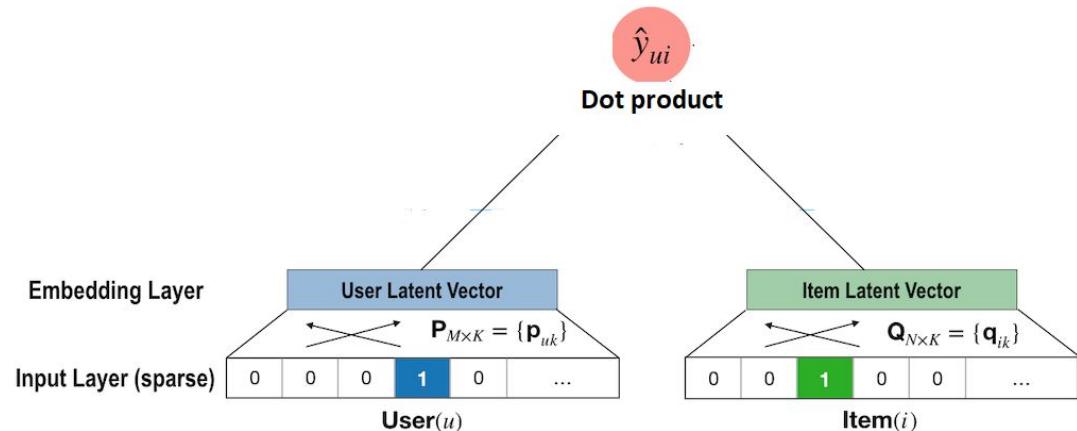
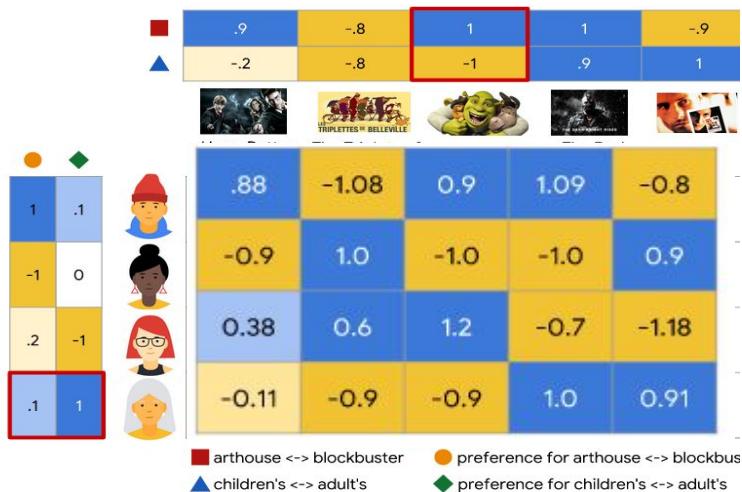
SVD

$$\begin{aligned} & |A - UV^T|_F^2 \\ &= \sum_{(i, j)} (A_{ij} - U_i \cdot V_j)^2 \end{aligned}$$

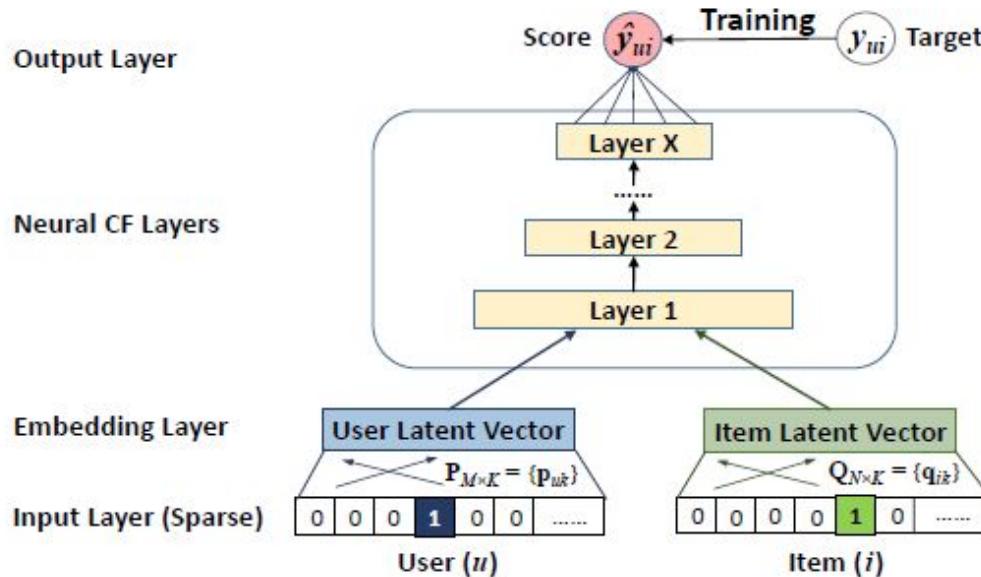
Weighted MF

$$\begin{aligned} & \sum_{(i, j) \in \text{obs}} (A_{ij} - U_i \cdot V_j)^2 + \\ & w_0 \sum_{(i, j) \notin \text{obs}} (0 - U_i \cdot V_j)^2 \end{aligned}$$

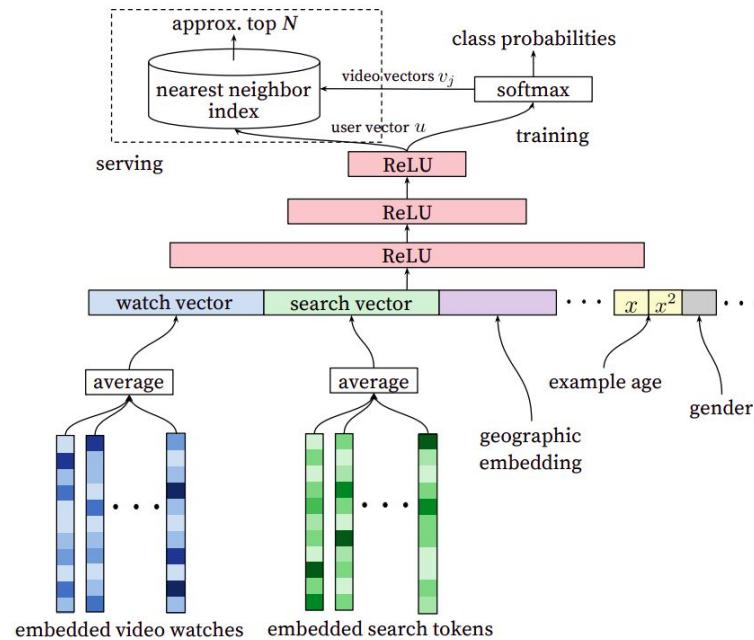
Neural Collaborative Filtering



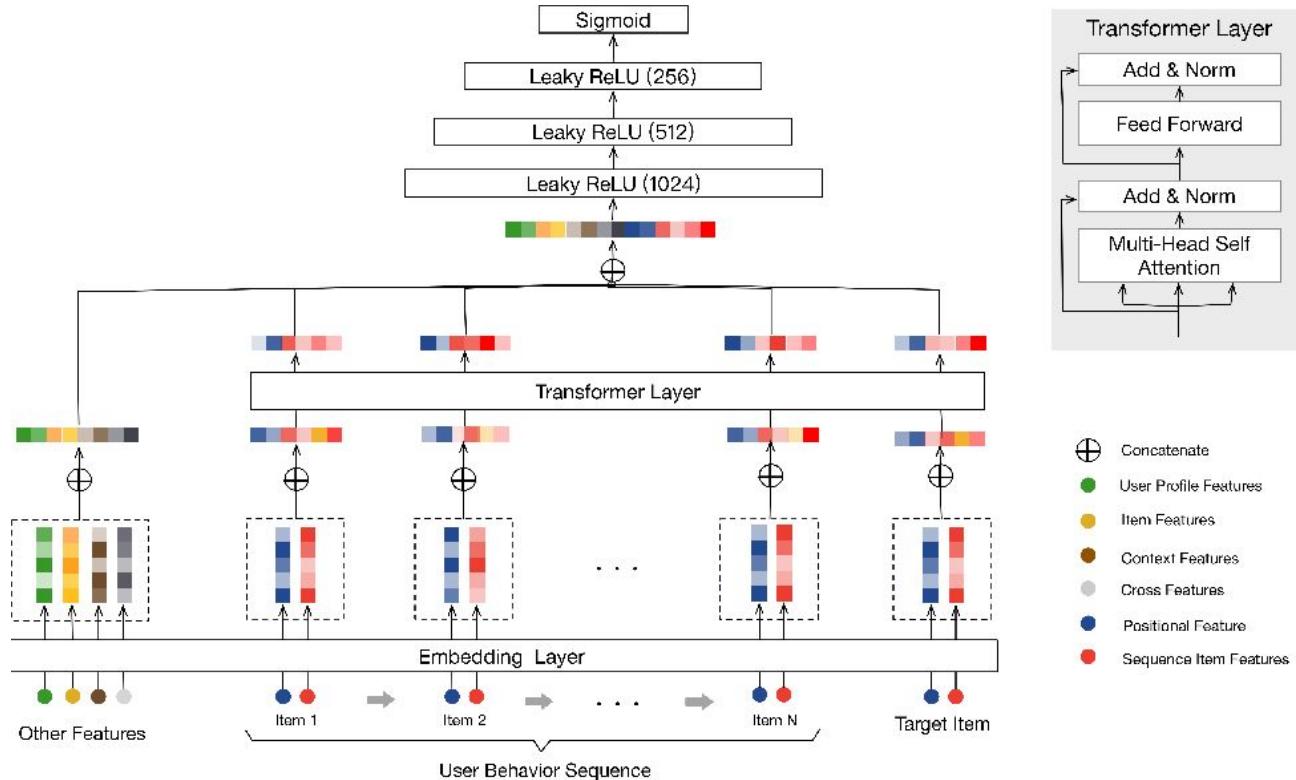
Neural Collaborative Filtering



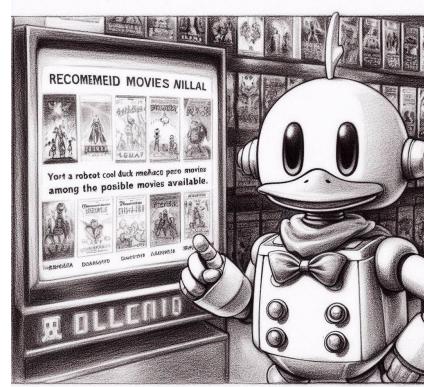
Neural Collaborative Filtering



Neural Collaborative Filtering



Introduction to recommender systems



AI frameworks

Evaluating recommender systems

Offline Evaluation

Evaluating the model before deployment:

- Using history of user-item interactions
- Choose a train/test split strategy: random-split, time-based split
- Filter users or items without enough interactions
- Choose a metric: explicit/implicit feedbacks

Offline Evaluation Metrics

					
u	4	5	4	3	5
					
	2.3	4.4	3.7	2.8	4.2

$$MAE = \frac{1}{N} \sum_{(u,i) \in \text{data}} |r_{ui} - \hat{r}_{ui}|$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{(u,i) \in \text{data}} (r_{ui} - \hat{r}_{ui})^2}$$

Explicit feedbacks:

- MAE
- RMSE

Offline Evaluation Metrics

u	0	1	0	1	1
	1	0	1	1	0

$$MAE = \frac{1}{N} \sum_{(u,i) \in \text{data}} |r_{ui} - \hat{r}_{ui}|$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{(u,i) \in \text{data}} (r_{ui} - \hat{r}_{ui})^2}$$

$$Precision = \frac{TP}{TP+FP}$$

Explicit feedbacks:

- MAE
- RMSE
- Precision
- Recall
- F1-Score

$$Recall = \frac{TP}{TP+FN}$$

$$F1_Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

Offline Evaluation Metrics



u



0.44 0.53 0.67 0.78

Implicit feedbacks:

Offline Evaluation Metrics



Mean Reciprocal Rank (MRR)

- Find first relevant recommendation rank k:
(here: $k = 2$)

$$MRR = \frac{1}{|U|} \sum_{u \in U} \frac{1}{k_u}$$

Implicit feedbacks:

- Mean Reciprocal Rank (MRR)

Offline Evaluation Metrics

Normalized Discounted Cumulative Gain (NDCG)



Implicit feedbacks:

- Mean Reciprocal Rank (MRR)
- Normalized Discounted Cumulative Gain (NDCG)

Offline Evaluation Metrics

Normalized Discounted Cumulative Gain (NDCG)



- Compute the cumulative gain:

$$DCG = \sum_{k=1}^K Gains = 0 + 1 + 0 + 1$$

Implicit feedbacks:

- Mean Reciprocal Rank (MRR)
- Normalized Discounted Cumulative Gain (NDCG)

Offline Evaluation Metrics



Normalized Discounted Cumulative Gain (NDCG)

- Compute the cumulative gain:

$$DCG = \sum_{k=1}^K Gains = 0 + 1 + 0 + 1$$

- Compute discounted cumulative gain (DCG):

$$\begin{aligned} DCG &= \sum_{k=1}^K \frac{Gains}{\log(k+1)} \\ &= \frac{0}{\log(1+1)} + \frac{1}{\log(2+1)} + \frac{0}{\log(3+1)} + \frac{1}{\log(4+1)} \end{aligned}$$

Implicit feedbacks:

- Mean Reciprocal Rank (MRR)
- Normalized Discounted Cumulative Gain (NDCG)

Offline Evaluation Metrics



Implicit feedbacks:

- Mean Reciprocal Rank (MRR)
- Normalized Discounted Cumulative Gain (NDCG)

Normalized Discounted Cumulative Gain (NDCG)

- Compute the cumulative gain:

$$DCG = \sum_{k=1}^K Gains = 0 + 1 + 0 + 1$$

- Compute discounted cumulative gain (DCG):

$$\begin{aligned} DCG &= \sum_{k=1}^K \frac{Gains}{\log(k+1)} \\ &= \frac{0}{\log(1+1)} + \frac{1}{\log(2+1)} + \frac{0}{\log(3+1)} + \frac{1}{\log(4+1)} \end{aligned}$$

- Compute Ideal Discounted Cumulative Gain (IDCG):

$$DCG = \frac{1}{\log(1+1)} + \frac{1}{\log(2+1)} + \frac{0}{\log(3+1)} + \frac{0}{\log(4+1)}$$

Offline Evaluation Metrics



Implicit feedbacks:

- Mean Reciprocal Rank (MRR)
- Normalized Discounted Cumulative Gain (NDCG)

Normalized Discounted Cumulative Gain (NDCG)

- Compute the cumulative gain:

$$DCG = \sum_{k=1}^K Gains = 0 + 1 + 0 + 1$$

- Compute discounted cumulative gain (DCG):

$$\begin{aligned} DCG &= \sum_{k=1}^K \frac{Gains}{\log(k+1)} \\ &= \frac{0}{\log(1+1)} + \frac{1}{\log(2+1)} + \frac{0}{\log(3+1)} + \frac{1}{\log(4+1)} \end{aligned}$$

- Compute Ideal Discounted Cumulative Gain (IDCG):

$$DCG = \frac{1}{\log(1+1)} + \frac{1}{\log(2+1)} + \frac{0}{\log(3+1)} + \frac{0}{\log(4+1)}$$

- Compute Normalized Discounted Cumulative Gain (NDCG):

$$NDCG = \frac{DCG}{IDCG}$$

Recommender systems in real-life

Real evaluation

Challenges of offline Evaluation:

- Biases in user interactions and ratings
- Highly sparse data

Real evaluation

Challenges of offline Evaluation:

- Bias in user interactions and ratings
- Highly sparse data

Online evaluation:

- Click-through rates
- Adoption and conversion (percentage of song listened, percentage of products bought, ...)
- Global revenue
- User behaviour and engagement (are the user coming more often? Do they stay longer?)
- A/B testing

Tricks

- Be careful to seasonality and behavioral change
- Be careful to bots
- Choose recommendation strategy depending on context
- Recommendation position is important
- Explaining why an item was recommended improves conversion rate
- Best sellers are an efficient option

Introduction to recommender systems



AI frameworks