

# Embeddings: Vector & Semantic Search

Applications | Concepts | Examples

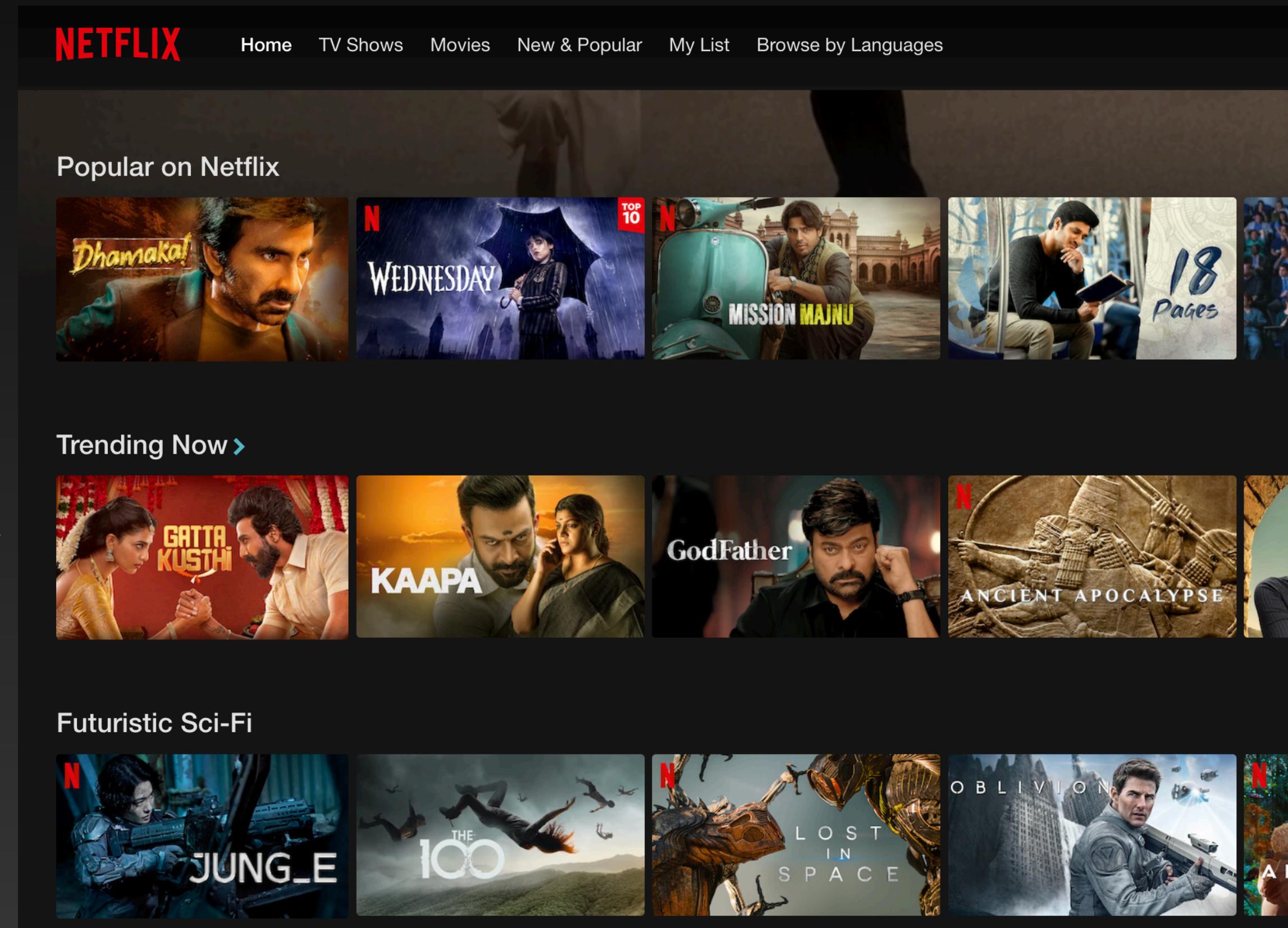
# Today's Lecture

**1. Introduction and Motivation**

**2. Semantic Search**

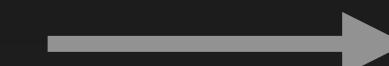
# Netflix Recommendations

MULTIPLE  
CAROUSELS

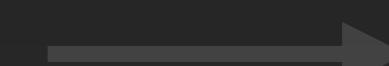


# Netflix Recommendations

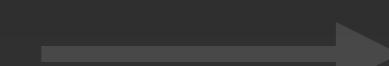
**POPULAR yet  
PERSONALIZED**



**PERSONALIZED  
TRENDS**



**PERSONALIZED  
GENRE**

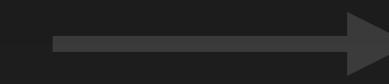


The screenshot shows the Netflix homepage with three main recommendation sections:

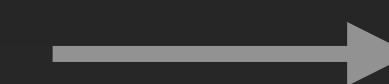
- Popular on Netflix:** Features movie posters for "Dhamaka!", "Wednesday" (marked as TOP 10), "MISSION MAJNU", "18 Pages", and "KAAPA".
- Trending Now >** Features movie posters for "GATTA KUSTHI", "KAAPA", "GodFather", "ANCIENT APOCALYPSE", and a partially visible poster.
- Futuristic Sci-Fi:** Features movie posters for "JUNG\_E", "THE 100", "LOST IN SPACE", "OBLIVION", and a partially visible poster.

# Netflix Recommendations

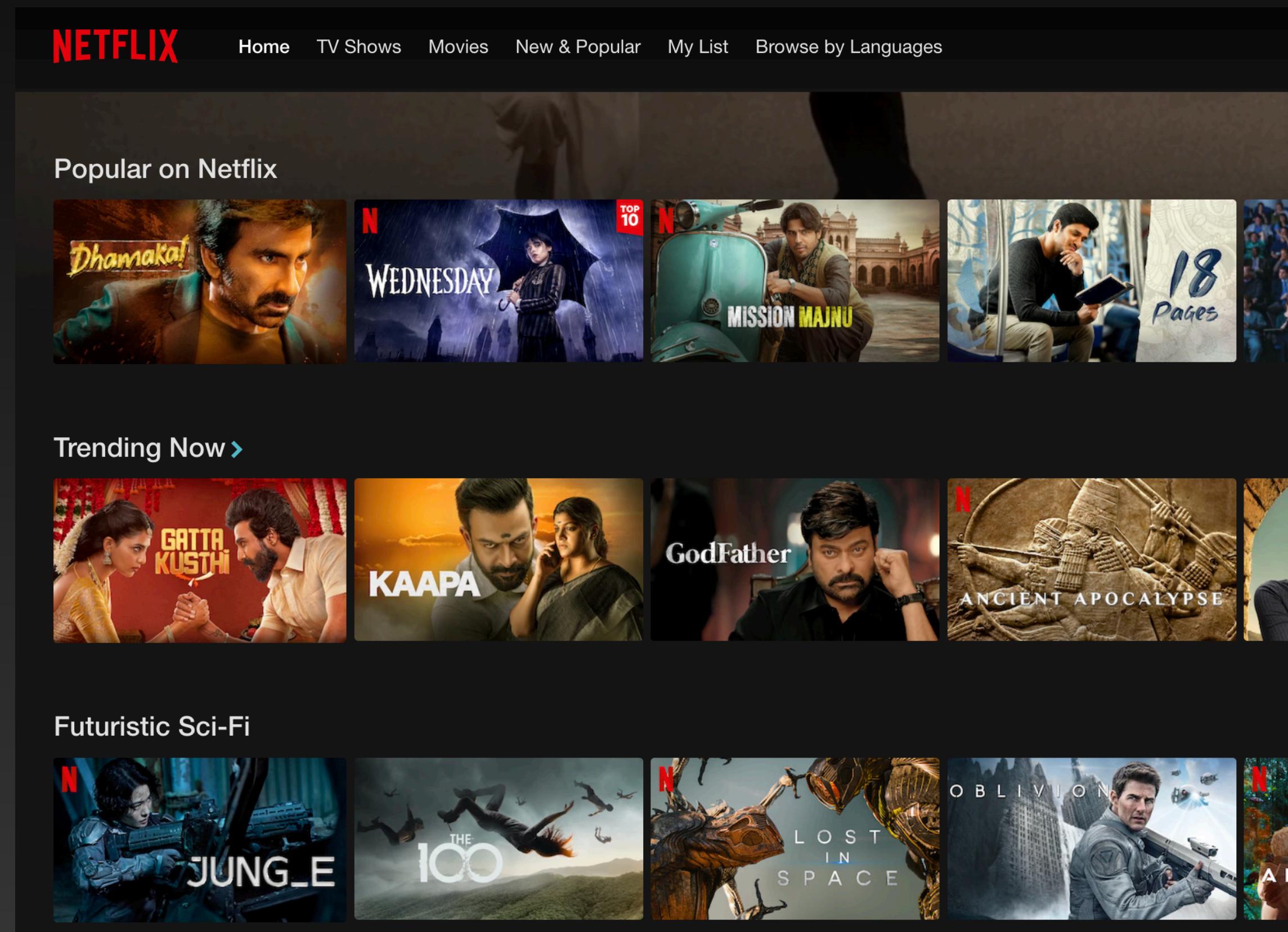
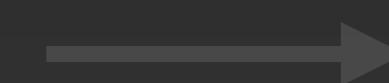
**POPULAR yet  
PERSONALIZED**



**PERSONALIZED  
TRENDS**

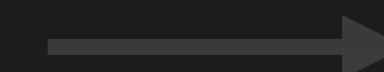


**PERSONALIZED  
GENRE**

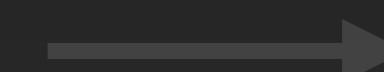


# Netflix Recommendations

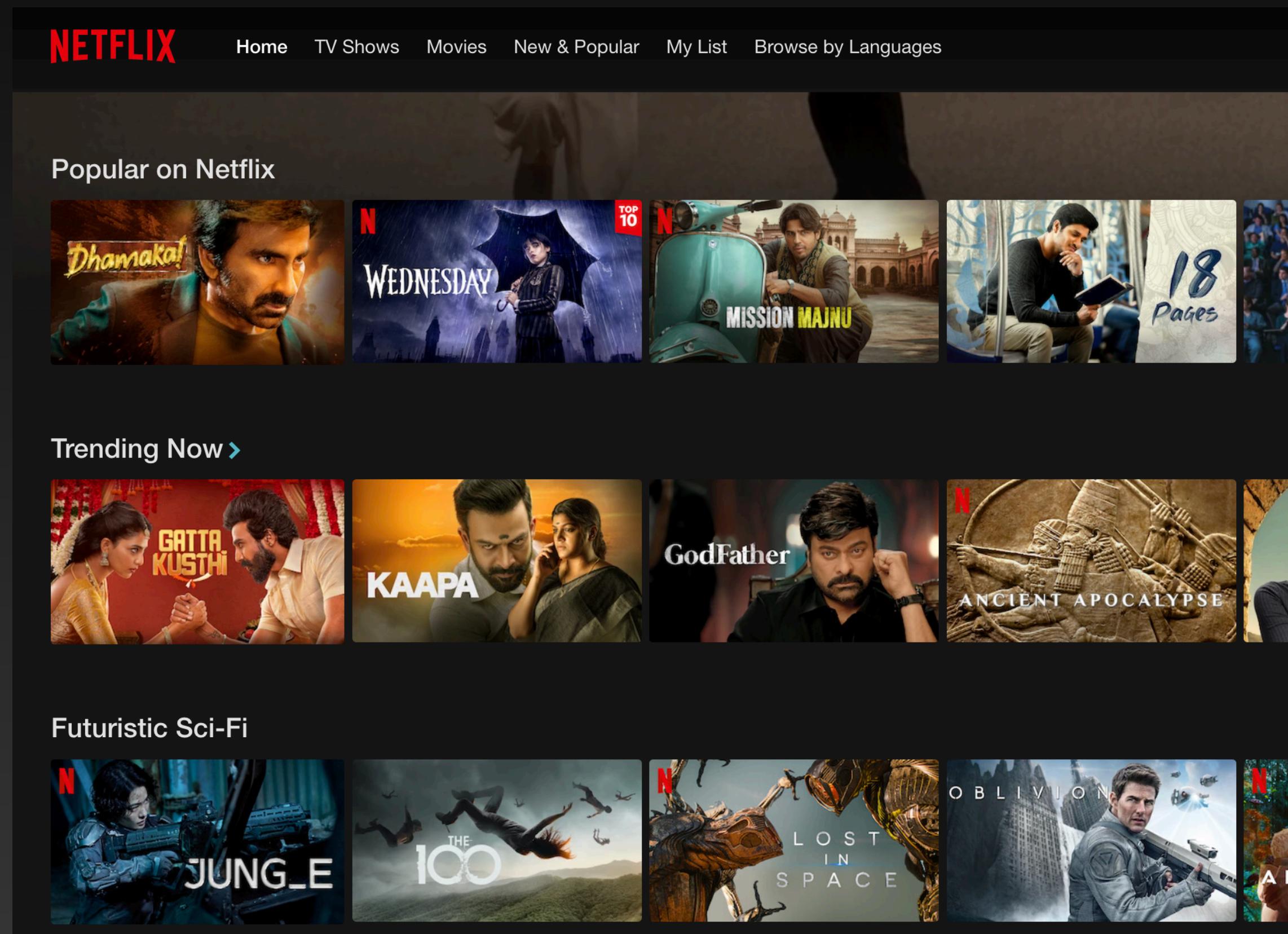
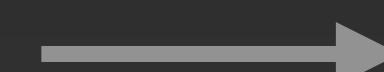
**POPULAR yet  
PERSONALIZED**



**PERSONALIZED  
TRENDS**



**PERSONALIZED  
GENRE**



# Netflix Million Dollar Prize!



# Netflix Million Dollar Prize!

The screenshot shows the Netflix Prize Leaderboard page. At the top, there's a red header bar with the Netflix logo and a yellow banner featuring three stars. Below the banner, the title "Netflix Prize" is displayed in large white letters. A navigation menu at the top includes links for Home, Rules, Leaderboard, Register, Update, Submit, and Download. The main content area has a blue header "Leaderboard" and a sub-header "Display top 40 leaders." The table below lists the top 40 teams, their scores, and submission times. The table has columns for Rank, Team Name, Best Score, % Improvement, Last Submit Time, and a header row.

Rank	Team Name	Best Score	% Improvement	Last Submit Time
-	No Grand Prize candidates yet	--	--	--
<b>Grand Prize - RMSE &lt;= 0.8563</b>				
1	<a href="#">PragmaticTheory</a>	0.8584	9.78	2009-06-16 01:04:47
2	<a href="#">BellKor in BigChaos</a>	0.8590	9.71	2009-05-13 08:14:09
3	<a href="#">Grand Prize Team</a>	0.8593	9.68	2009-06-12 08:20:24
4	<a href="#">Dace</a>	0.8604	9.56	2009-04-22 05:57:03
5	<a href="#">BigChaos</a>	0.8613	9.47	2009-06-15 18:03:55
<b>Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos</b>				
6	<a href="#">BellKor</a>	0.8620	9.40	2009-06-17 13:41:48
7	<a href="#">Gravity</a>	0.8634	9.25	2009-04-22 18:31:32
8	<a href="#">Opera Solutions</a>	0.8640	9.19	2009-06-09 22:24:53
9	<a href="#">xlivector</a>	0.8640	9.19	2009-06-17 12:47:27
10	<a href="#">BruceDengDaoCiYiYou</a>	0.8641	9.18	2009-06-02 17:08:31
11	<a href="#">Ces</a>	0.8642	9.17	2009-06-12 23:04:25
12	<a href="#">majia2</a>	0.8642	9.17	2009-06-15 03:35:00
13	<a href="#">xiangliang</a>	0.8642	9.17	2009-06-13 16:35:35
14	<a href="#">Feeds2</a>	0.8647	9.11	2009-06-16 22:21:19
15	<a href="#">Just a guy in a garage</a>	0.8650	9.08	2009-05-24 10:02:54
16	<a href="#">Team ESP</a>	0.8653	9.05	2009-06-16 05:25:11
17	<a href="#">pengpengzhou</a>	0.8654	9.04	2009-05-05 18:18:03
18	<a href="#">NewNetflixTeam</a>	0.8657	9.01	2009-05-31 07:30:22
19	<a href="#">J Dennis Su</a>	0.8658	9.00	2009-03-11 09:41:54

# Collaborative Filtering

**Rishi**

**Michael**

**Karthik**

**Roshin**

**Amy**

# Collaborative Filtering



Avatar



Arrival



When Harry



Before Sunrise



Minions

**Rishi**

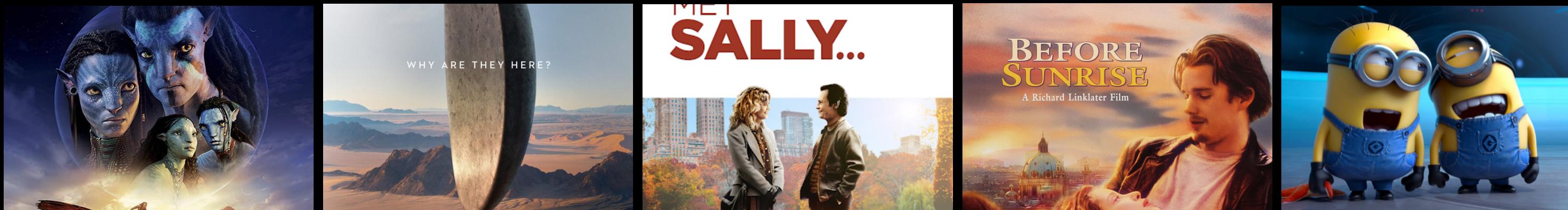
**Michael**

**Karthik**

**Roshin**

**Amy**

# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



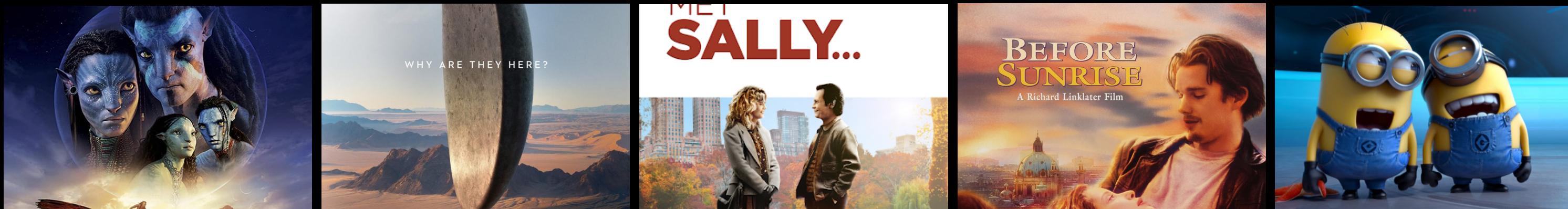
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



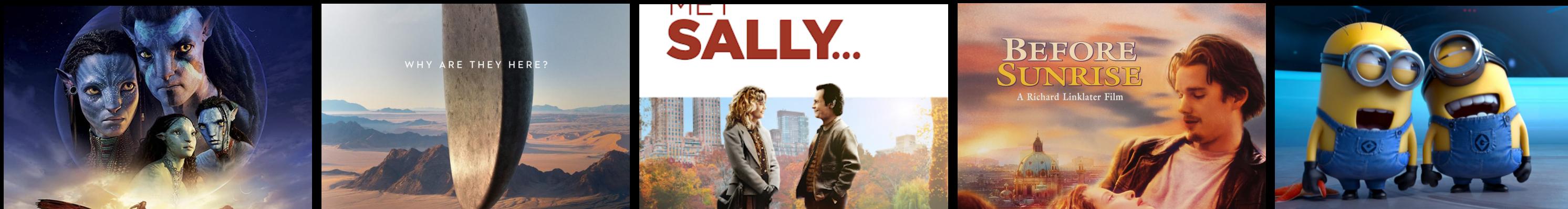
Roshin



Amy



# Collaborative Filtering



Rishi

Michael

Karthik

Roshin

Amy

	Avatar	Arrival	When Harry	Before Sunrise	Minions
Rishi	👍	👍	👎	👎	👍
Michael	👎	?	👍	👍	👍
Karthik	👍	?	👎	?	👍
Roshin	?	👎	👍	👍	?
Amy	?	👍	?	👍	👎

# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

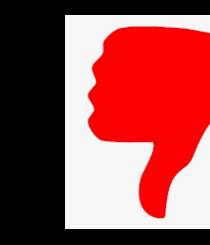
Rishi



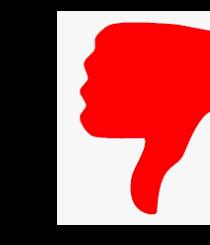
Michael



Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

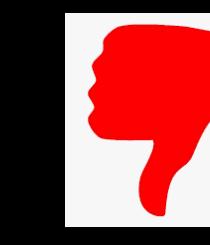
Arrival

When Harry

Before Sunrise

Minions

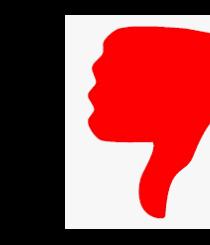
Rishi



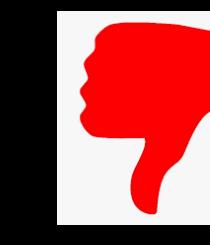
Michael



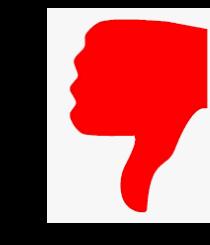
Karthik



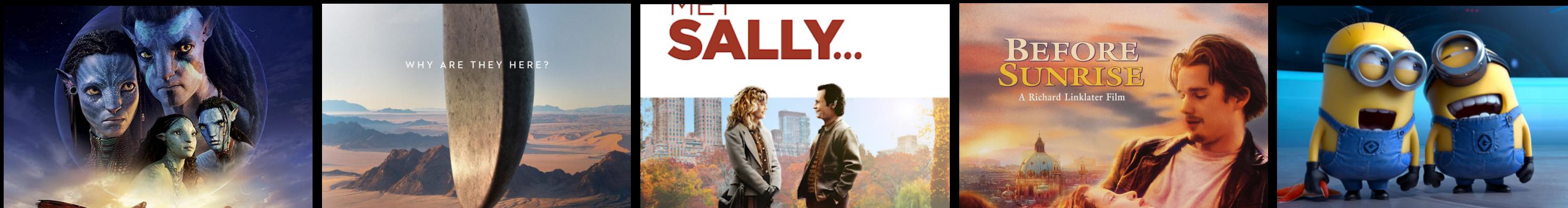
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



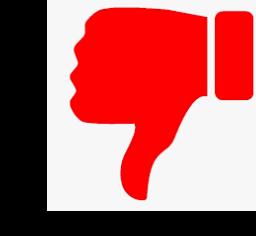
Karthik



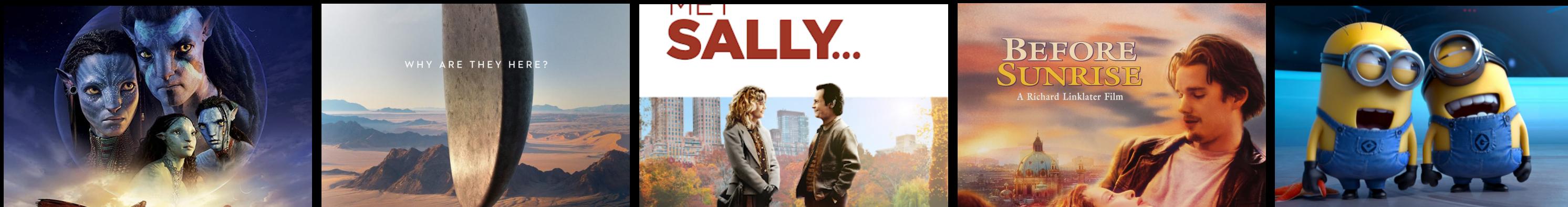
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

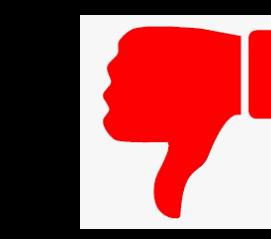
Rishi



Michael



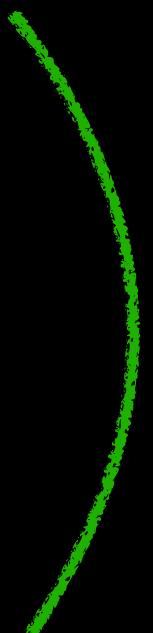
Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

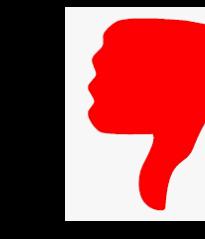
Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



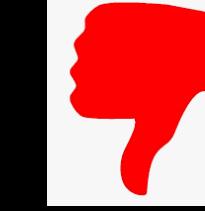
Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



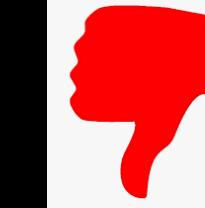
Karthik



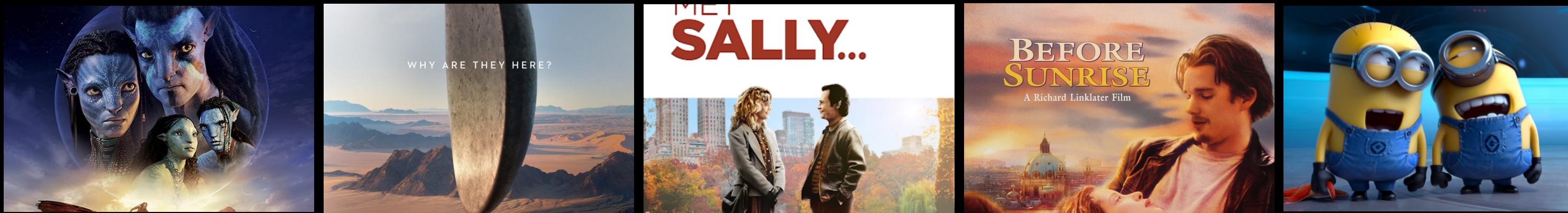
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

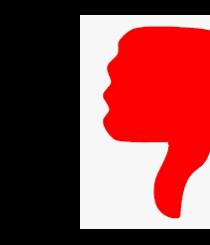
Rishi



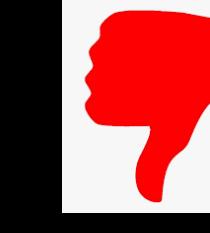
Michael



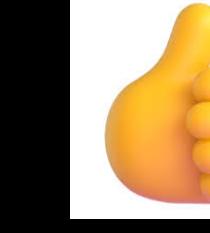
Karthik



Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



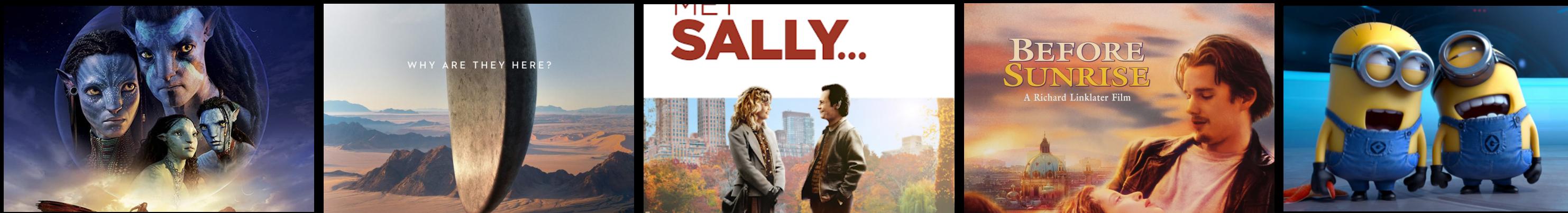
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



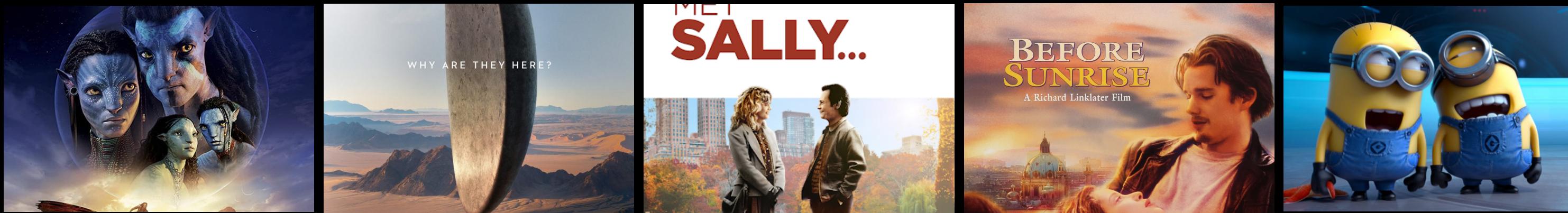
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



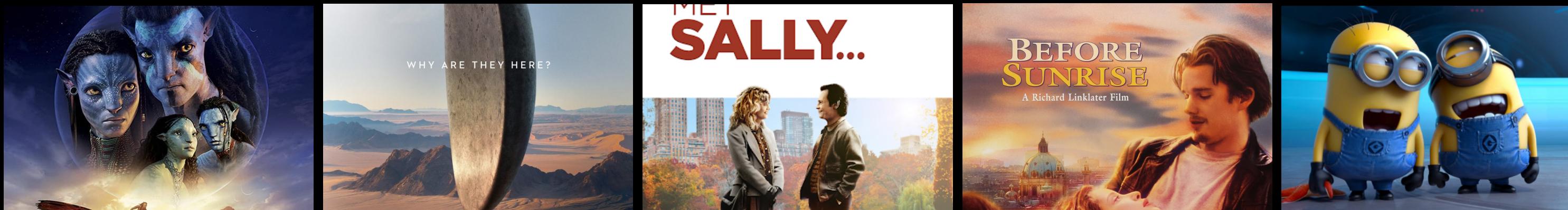
Roshin



Amy



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Rishi



Michael



Karthik



Roshin



Amy



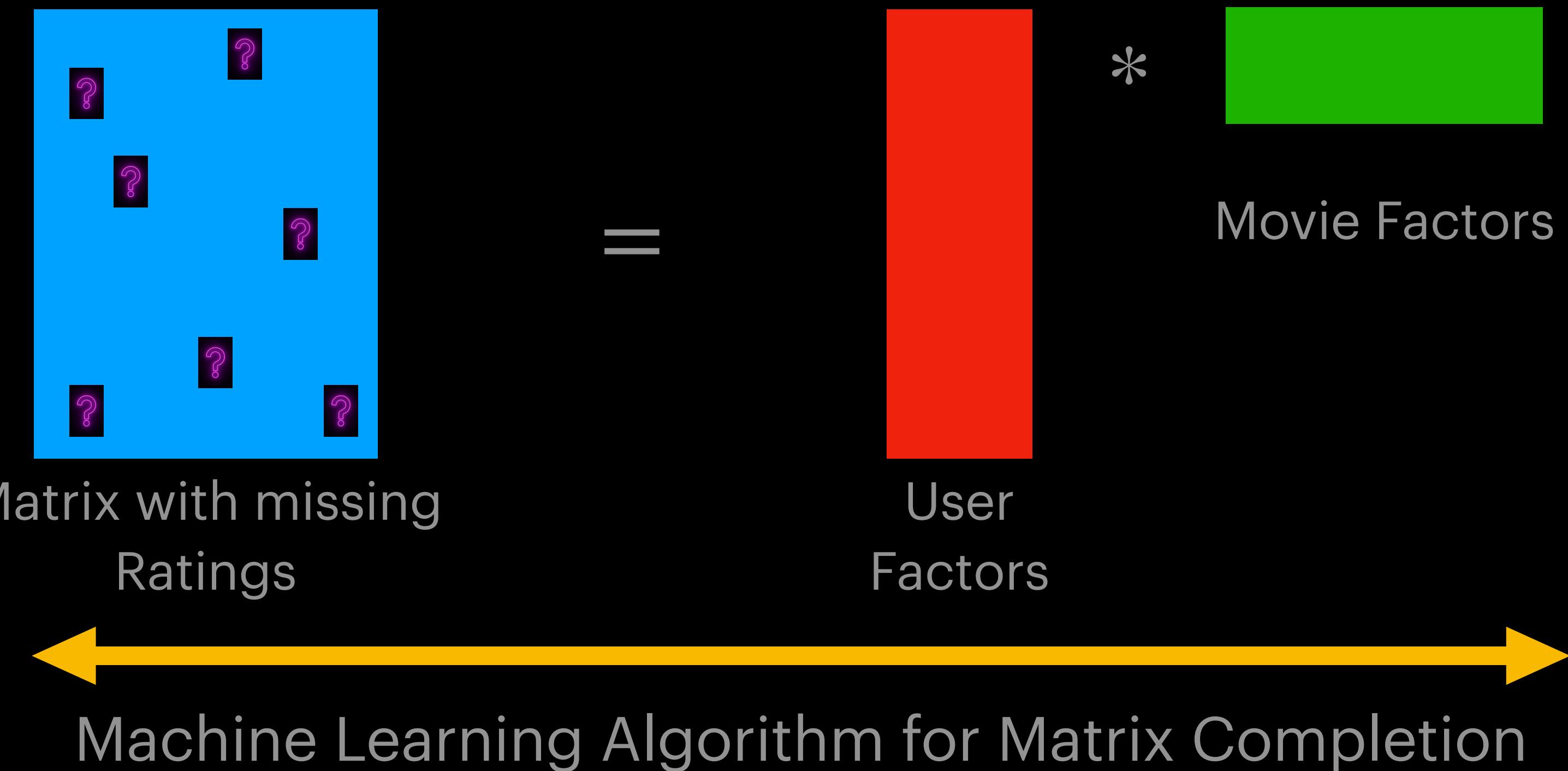
# Collaborative Filtering

## Through Matrix Completion!

👍	👍	👎	👎	👍
👎	?	👍	👍	👍
👍	?	👎	?	👍
?	👎	👍	👍	?
?	👍	?	👍	👎

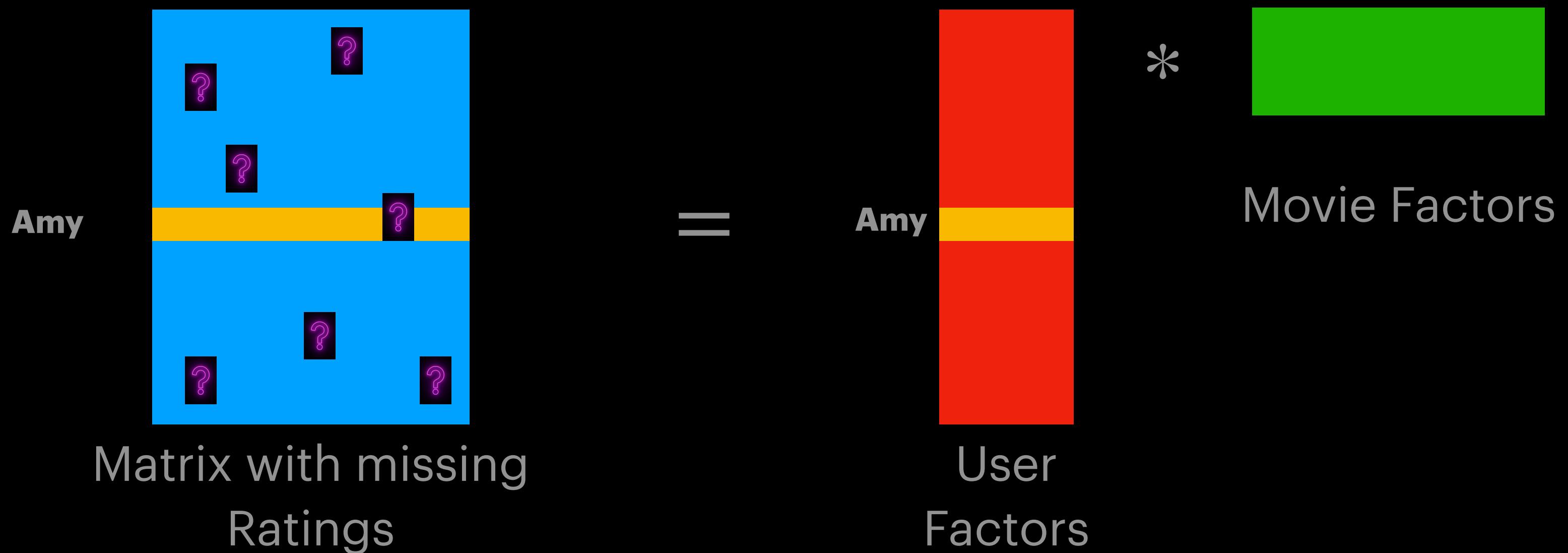
# Collaborative Filtering

## Through Matrix Completion!



# Collaborative Filtering

## Through Matrix Completion!



# Collaborative Filtering

## Through Matrix Completion!



# Collaborative Filtering

## Through Matrix Completion!



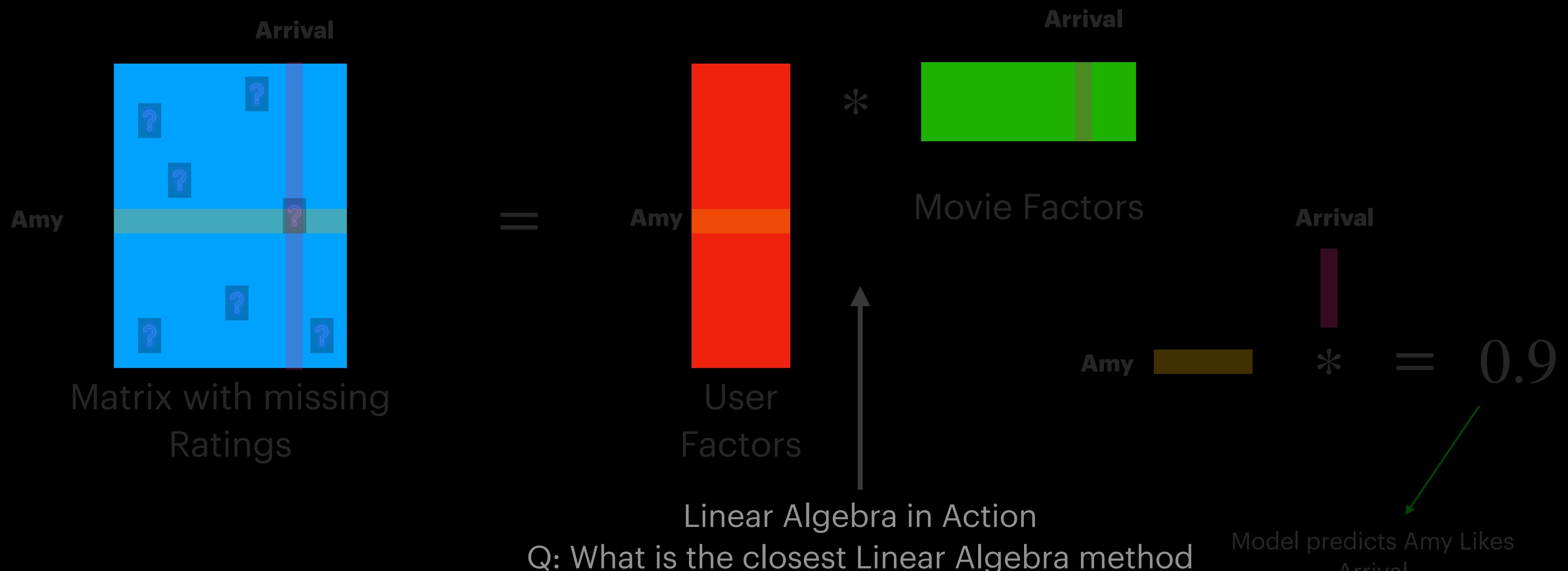
# Collaborative Filtering

## Through Matrix Completion!



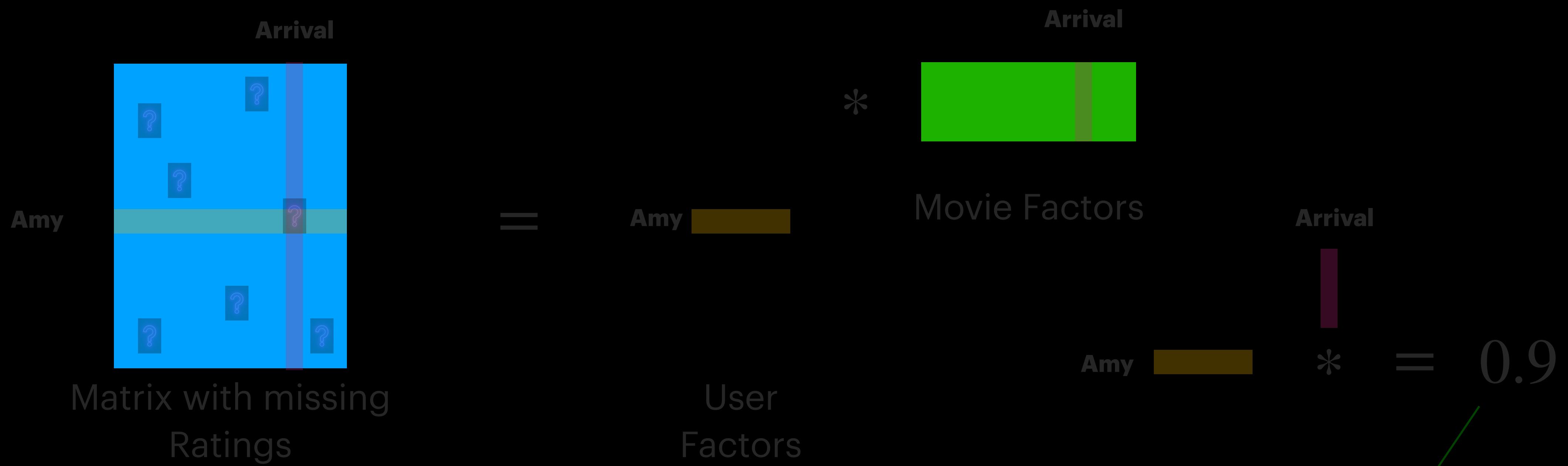
# Collaborative Filtering

## Through Matrix Completion!



# Collaborative Filtering

## Through Matrix Completion!



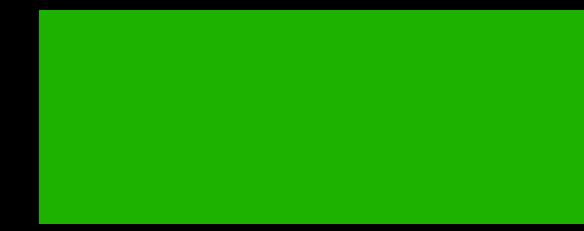
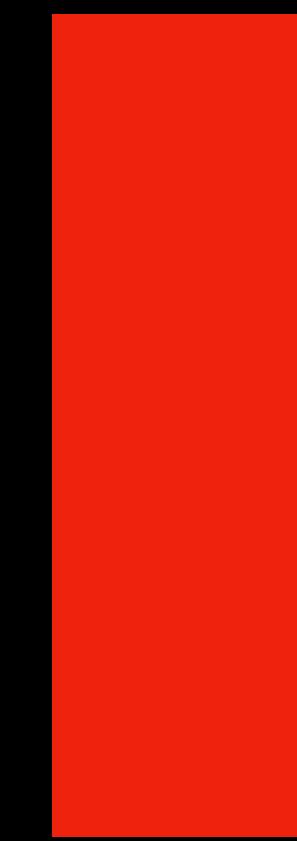
Linear Algebra in Action

Q: What is the closest Linear Algebra method that looks similar to the above factorization?

**A: SVD = Singular Value Decomposition**



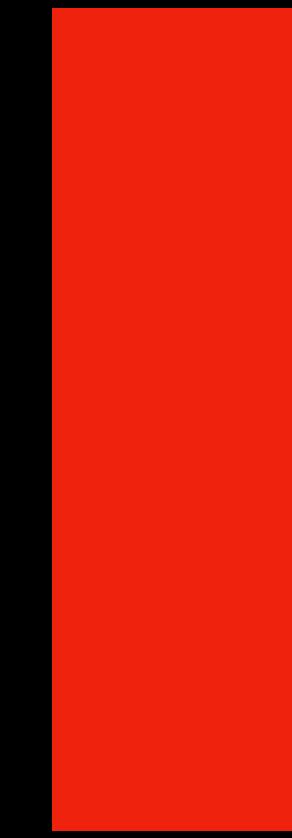
# SVD of a matrix



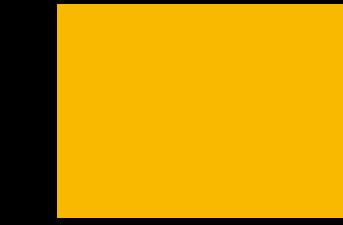
# SVD of a matrix



x



U



Sigma

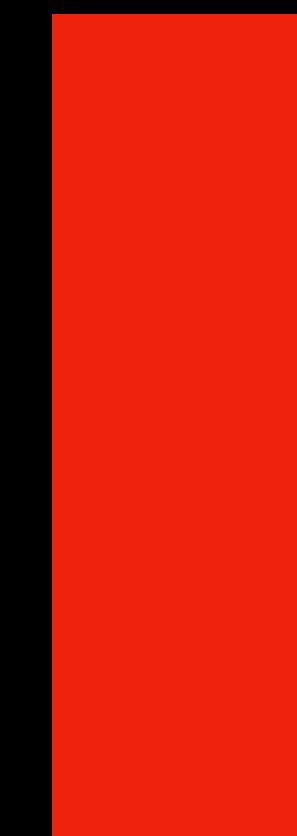


V^T

# SVD of a matrix



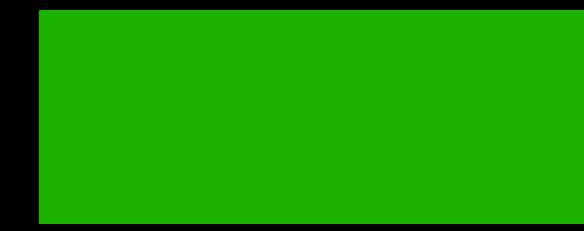
x



U



Sigma

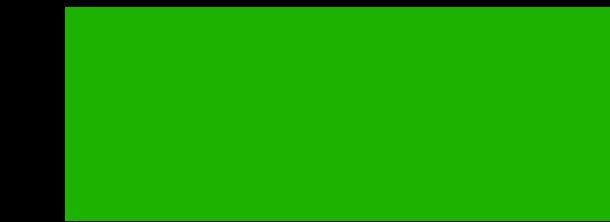
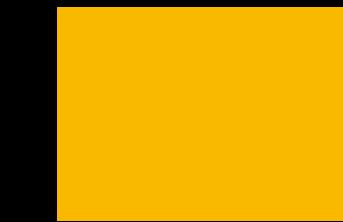
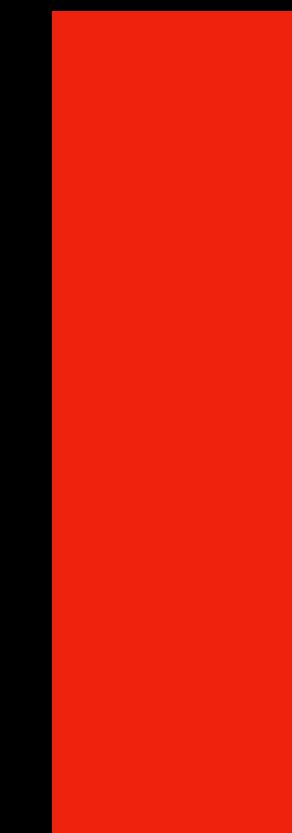


V^T

$$X = U\Sigma V^T$$

# SVD of a matrix

**Every matrix has an SVD!**



x

U

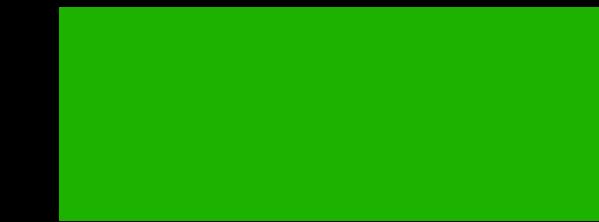
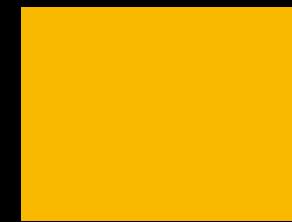
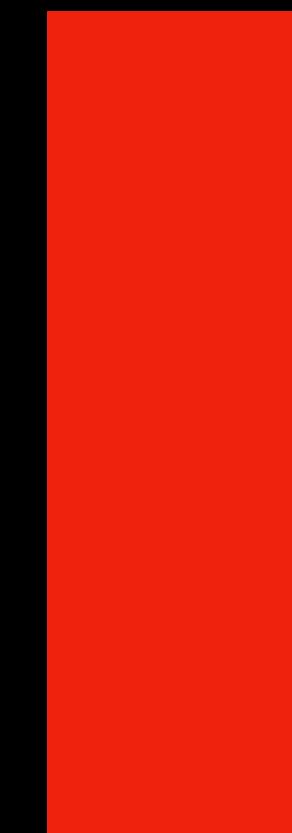
Sigma

V^T

$$X = U\Sigma V^T$$

# SVD of a matrix

Hence: Data Matrix also has an SVD!



X

U

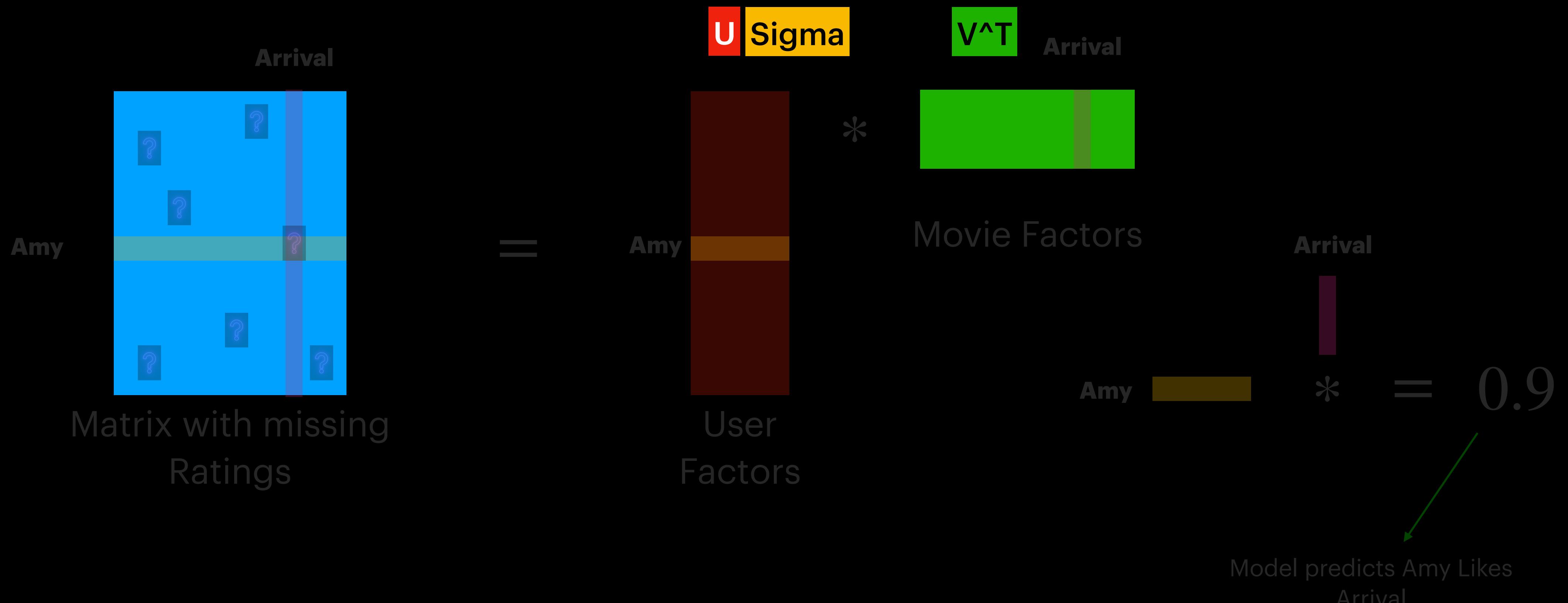
Sigma

V^T

$$X = U\Sigma V^T$$

# Collaborative Filtering

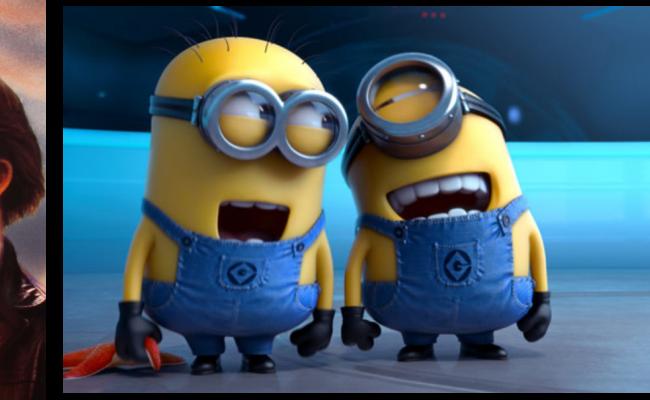
## Through Matrix Completion!



Collaborative Filtering: Advanced SVD method or Iterative SVD method



# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Men in Black

	Avatar	Arrival	When Harry	Before Sunrise	Minions	Men in Black
U1	👍	👍	👎	👎	👍	
U2	👎	👎	👍	👍	👍	
U3	👍	👍	👎	👎	👍	
U4	👎	👎	👍	👍	👍	
U5	👍	👍	👍	👍	👎	

# Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Men in Black

	Avatar	Arrival	When Harry	Before Sunrise	Minions	Men in Black
U1	👍	👍	👎	👎	👍	?
U2	👎	👎	👍	👍	👍	?
U3	👍	👍	👎	👎	👍	?
U4	👎	👎	👍	👍	👍	?
U5	👍	👍	👍	👍	👎	?

# Collaborative Filtering



Avatar

Arrival

When Harry

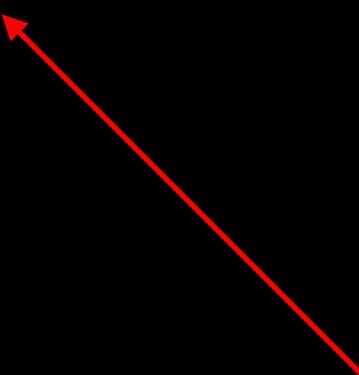
Before Sunrise

Minions

Men in Black

	Avatar	Arrival	When Harry	Before Sunrise	Minions	Men in Black
U1	👍	👍	👎	👎	👍	?
U2	👎	👎	👍	👍	👍	?
U3	👍	👍	👎	👎	👍	?
U4	👎	👎	👍	👍	👍	?
U5	👍	👍	👍	👍	👎	?

Cold Start Problem



# Content Based Filtering vs Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

Men in Black

	Avatar	Arrival	When Harry	Before Sunrise	Minions	
U1	👍	👍	👎	👎	👍	
U2	👎	👎	👍	👍	👍	
U3	👍	👍	👎	👎	👍	
U4	👎	👎	👍	👍	👍	
U5	👍	👍	👍	👍	👎	

Karthik

# Content Based Filtering vs Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

	Avatar	Arrival	When Harry	Before Sunrise	Minions	
U1	👍	👍	👎	👎	👍	
U2	👎	👎	👍	👍	👍	Karthik
U3	👍	👍	👎	👎	👍	
U4	👎	👎	👍	👍	👍	
U5	👍	👍	👍	👍	👎	

# Content Based Filtering vs Collaborative Filtering



Avatar

Arrival

When Harry

Before Sunrise

Minions

	Avatar	Arrival	When Harry	Before Sunrise	Minions
U1	👍	👍	👎	👎	👍
U2	👎	👎	👍	👍	👍
U3	👍	👍	👎	👎	👍
U4	👎	👎	👍	👍	👍
U5	👍	👍	👍	👍	👎



Watched



Arrival

Karthik

# Content Based Filtering vs Collaborative Filtering



Avatar

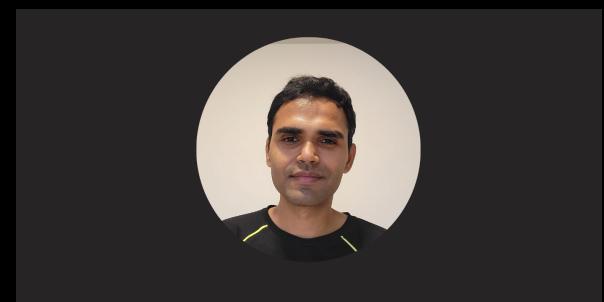
Arrival

When Harry

Before Sunrise

Minions

	Avatar	Arrival	When Harry	Before Sunrise	Minions
U1	👍	👍	👎	👎	👍
U2	👎	👎	👍	👍	👍
U3	👍	👍	👎	👎	👍
U4	👎	👎	👍	👍	👍
U5	👍	👍	👍	👍	👎



Watched



Karthik



# Content Based Filtering vs Collaborative Filtering



Avatar

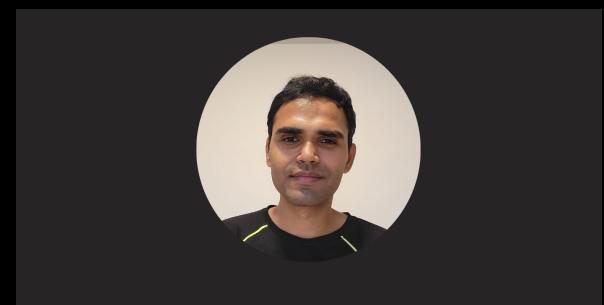
Arrival

When Harry

Before Sunrise

Minions

	Avatar	Arrival	When Harry	Before Sunrise	Minions
U1	👍	👍	👎	👎	👍
U2	👎	👎	👍	👍	👍
U3	👍	👍	👎	👎	👍
U4	👎	👎	👍	👍	👍
U5	👍	👍	👍	👍	👎



Watched



Arrival

Karthik

Likely  
To  
Watch



Men in Black

# Content Based Filtering



Minions



Men in Black



Men in Black



Arrival



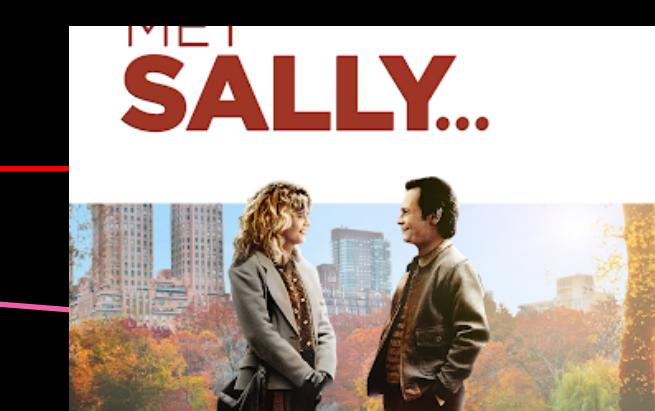
Arrival



When Harry met Sally



When Harry met Sally



When Harry met Sally



# Content Based Filtering

## Embeddings



Men in Black



Arrival



When Harry met



Typically 128 or 256  
latent dimensions

# Content Based Filtering

## Embeddings



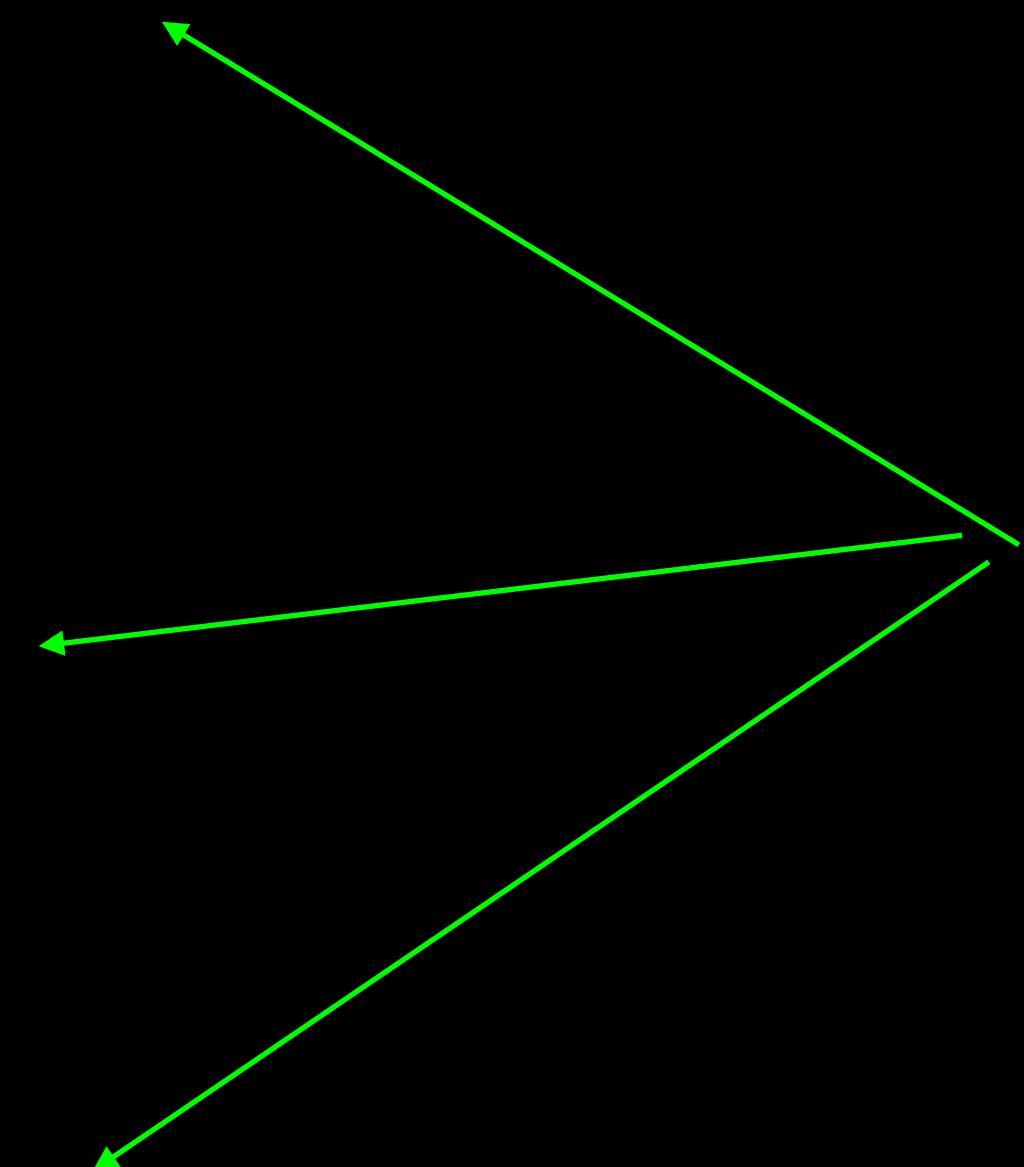
Men in Black



Arrival



When Harry met



**How do we  
Obtain these embeddings?**

# Content Based Filtering

## Embeddings



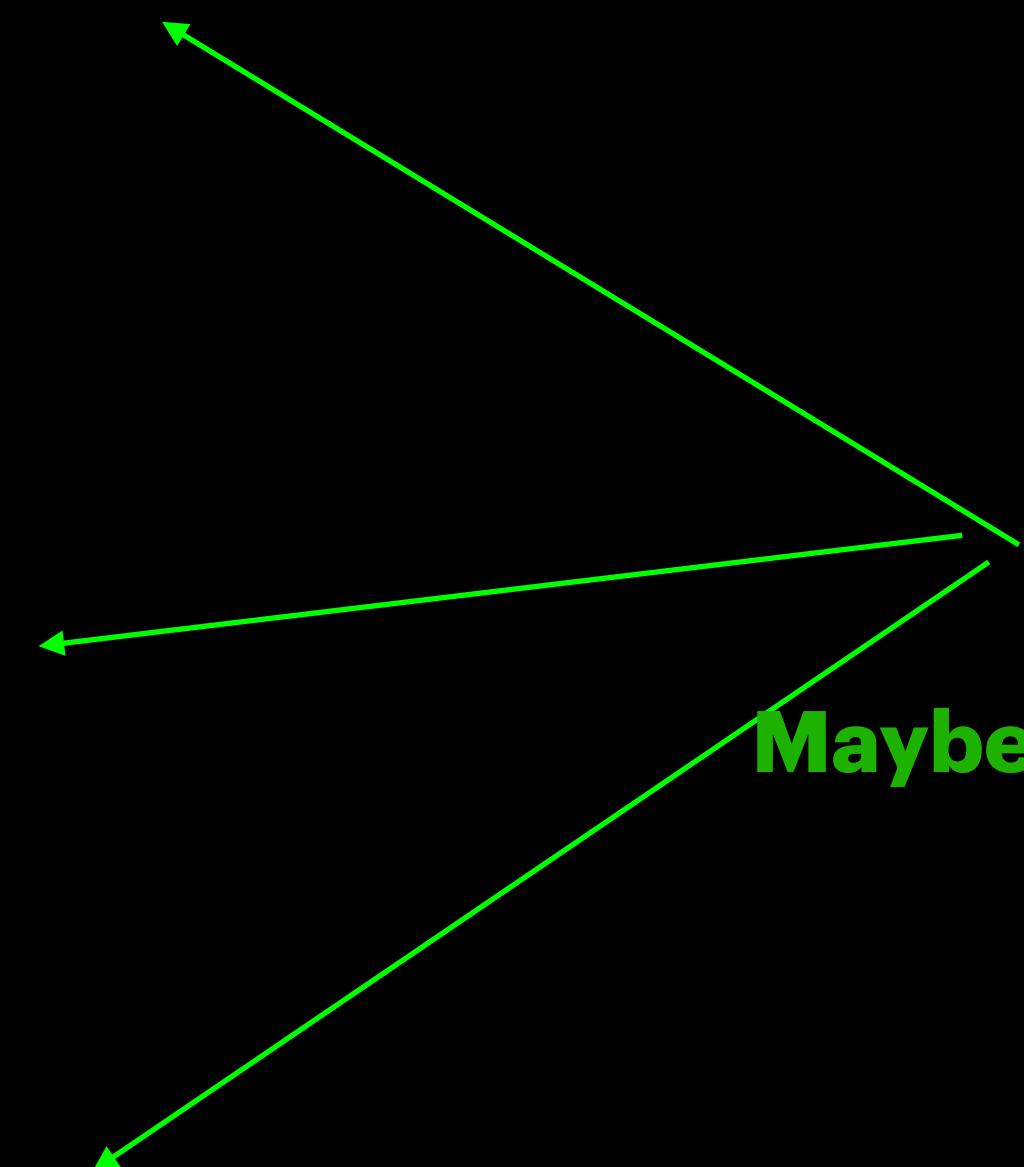
Men in Black



Arrival



When Harry met



**How do we  
Obtain these embeddings?**

**A: Through a DL model!  
Maybe last but one hidden layer activations**

# Embeddings Interpretation

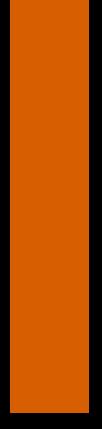
Embeddings



Men in Black



Arrival



When Harry met

Latent Dimensions

→ Interpretation

?      ?      ?      ?  
⋮      ⋮      ⋮      ⋮

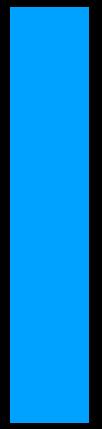
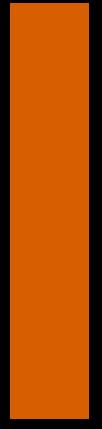
- Comedy + Romance
- Thriller + Adventure
- Comedy + Thriller
- Adventure
- ⋮      ⋮

# Embeddings | Vector Representations

Embeddings



Men in Black



Arrival



When Harry met

Latent Dimensions → Interpretation

?      ?      ?      ?  
⋮      ⋮      ⋮      ⋮

- Comedy + Romance
- Thriller + Adventure
- Comedy + Thriller
- Adventure
- ⋮      ⋮

# Embeddings in Action



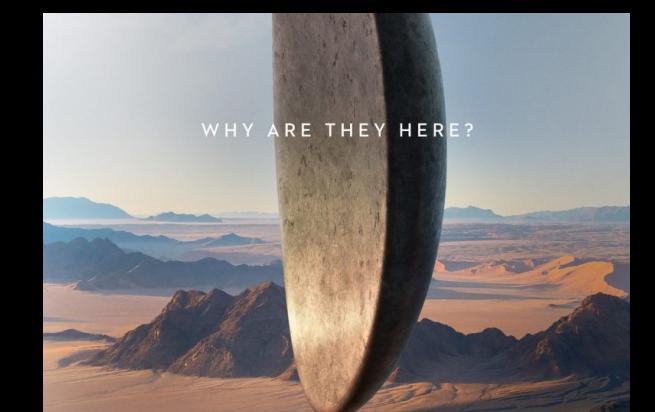
Minions



Men in Black



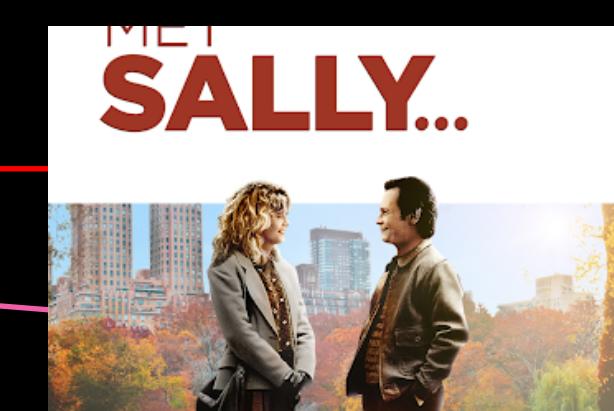
Men in Black



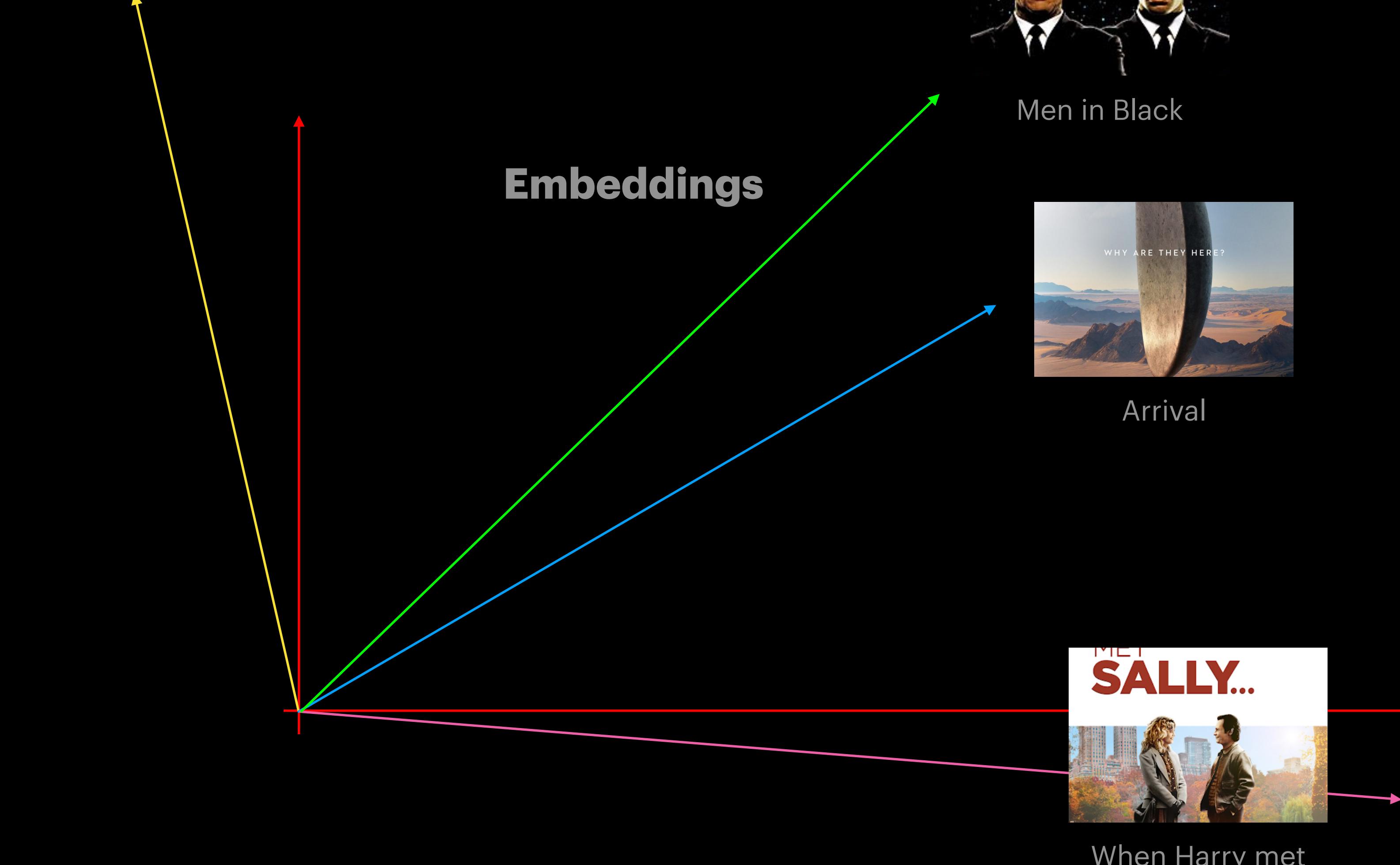
Arrival



When Harry met



When Harry met



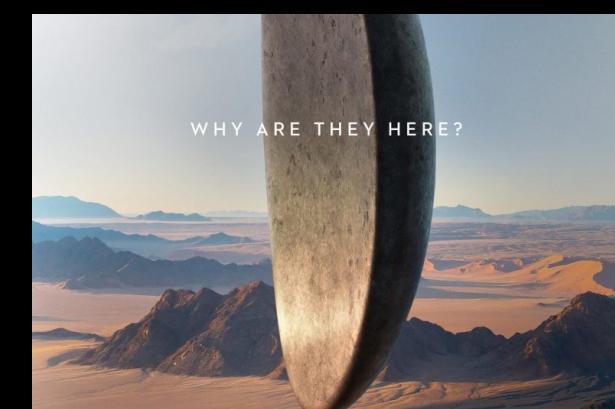
# Embeddings in Action



Minions



Men in Black



Arrival



Arrival



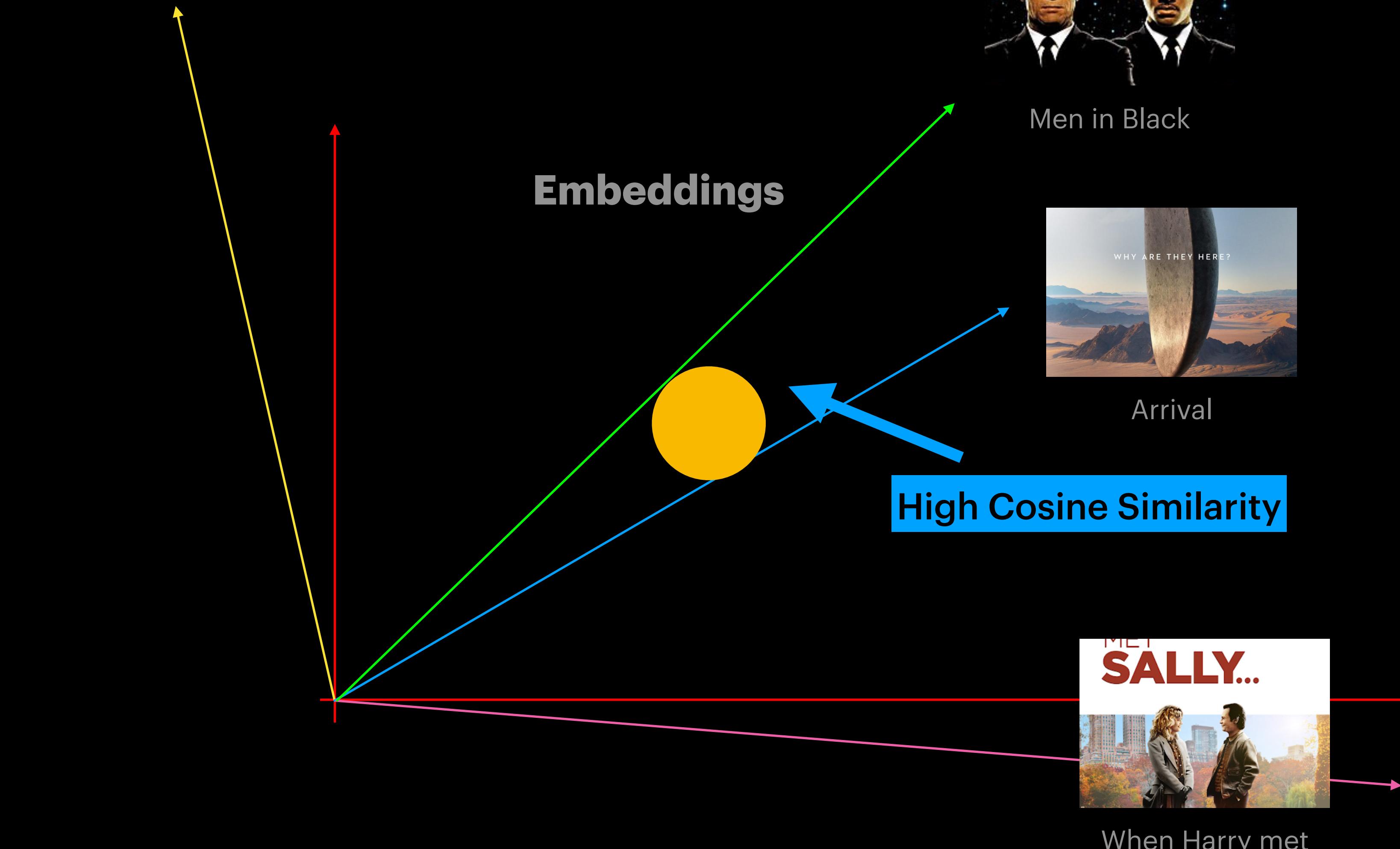
When Harry met



When Harry met

Embeddings

High Cosine Similarity



# Embeddings in Action



Minions



Men in Black



Arrival



Arrival



When Harry met

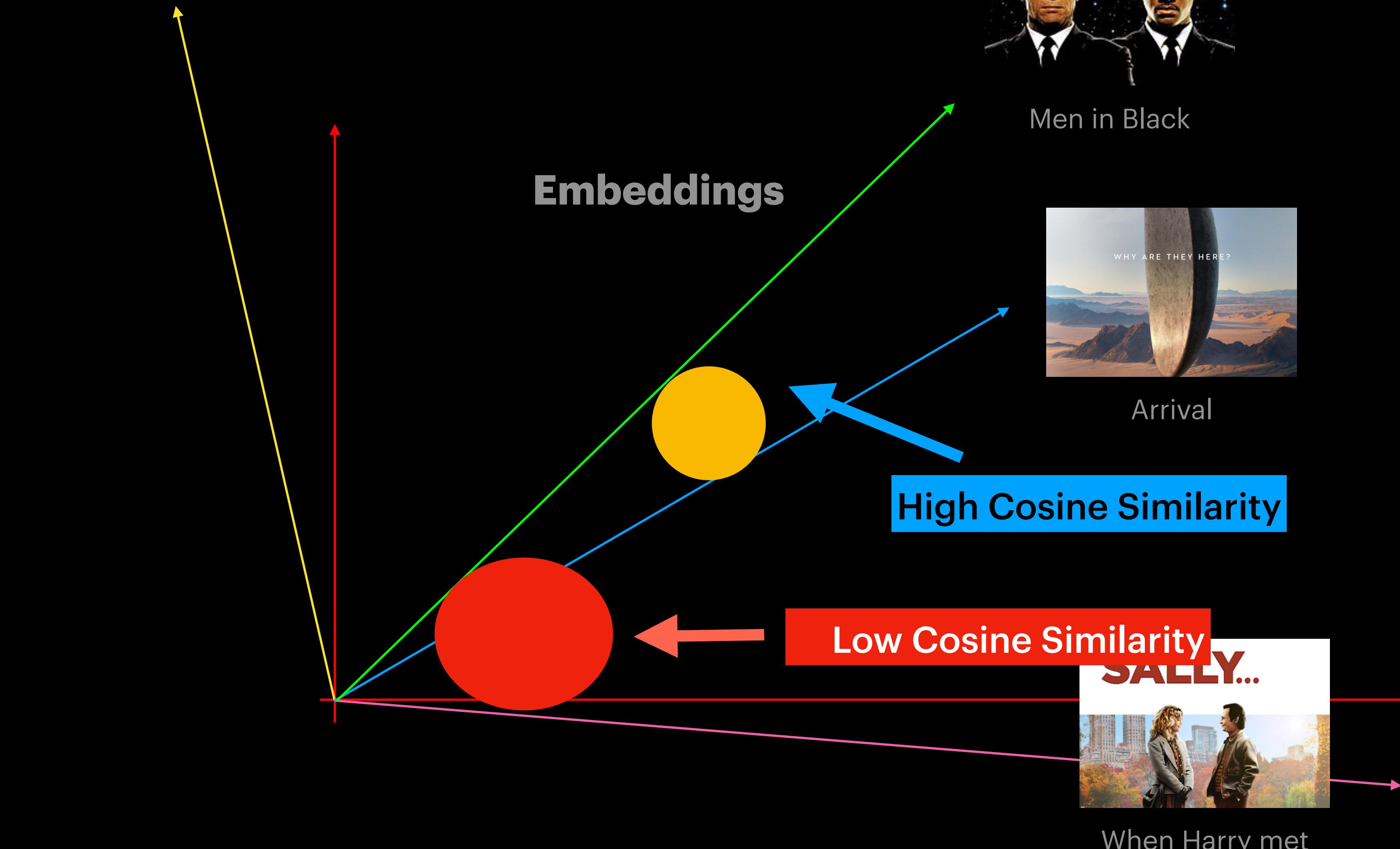


When Harry met

Embeddings

High Cosine Similarity

Low Cosine Similarity



# Embeddings in Action



Minions



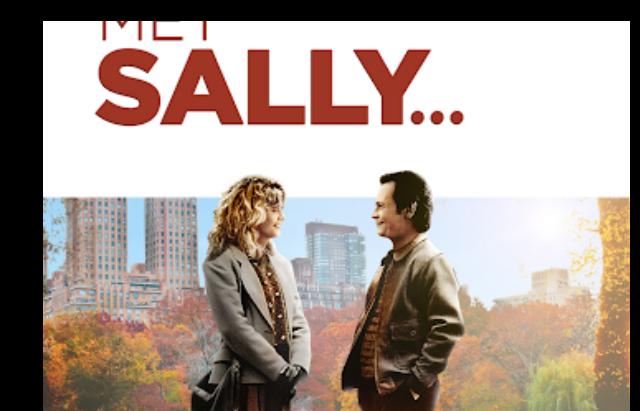
Men in Black



Men in Black



Arrival

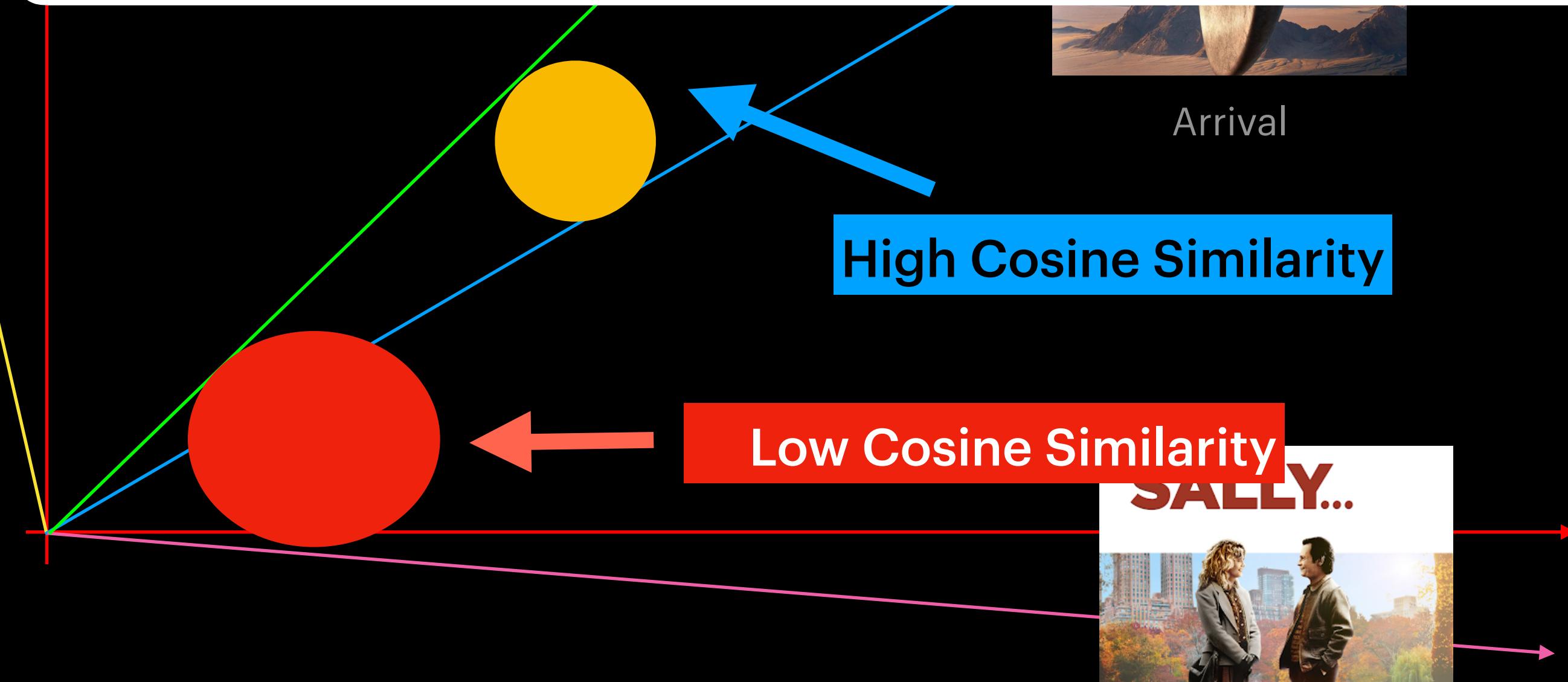


When Harry met



When Harry met

**Smaller Angle = Higher Cosine Similarity**  
**Larger Angle = Lower Cosine Similarity**



# Embeddings in Action



Minions



Men in Black



Men in Black



Arrival



When Harry met



Embeddings

**Smaller Angle = Higher Cosine Similarity**

**Larger Angle = Lower Cosine Similarity**

High Cosine Similarity

Low Cosine Similarity

$$? \leq \text{CosineSimilarity}(x, y) = \frac{x^T y}{\|x\| \|y\|} \leq ?$$

# Embeddings in Action



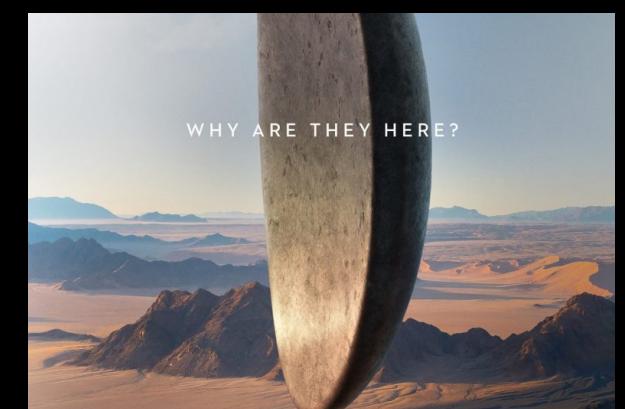
Minions



Men in Black



Men in Black



Arrival



When Harry met



Embeddings

**Smaller Angle = Higher Cosine Similarity**

**Larger Angle = Lower Cosine Similarity**

High Cosine Similarity

Low Cosine Similarity

$$-1 \leq \text{CosineSimilarity}(x, y) = \frac{x^T y}{\|x\| \|y\|} \leq 1$$

# Cauchy-Schwarz Inequality!

$$-1 \leq \text{CosineSimilarity}(x, y) = \frac{x^T y}{\|x\| \|y\|} \leq 1$$

# Cauchy-Schwarz Inequality!

$$-1 \leq \text{CosineSimilarity}(x, y) = \frac{x^T y}{\|x\| \|y\|} \leq 1$$

$$|x^T y| \leq \|x\| \|y\|$$

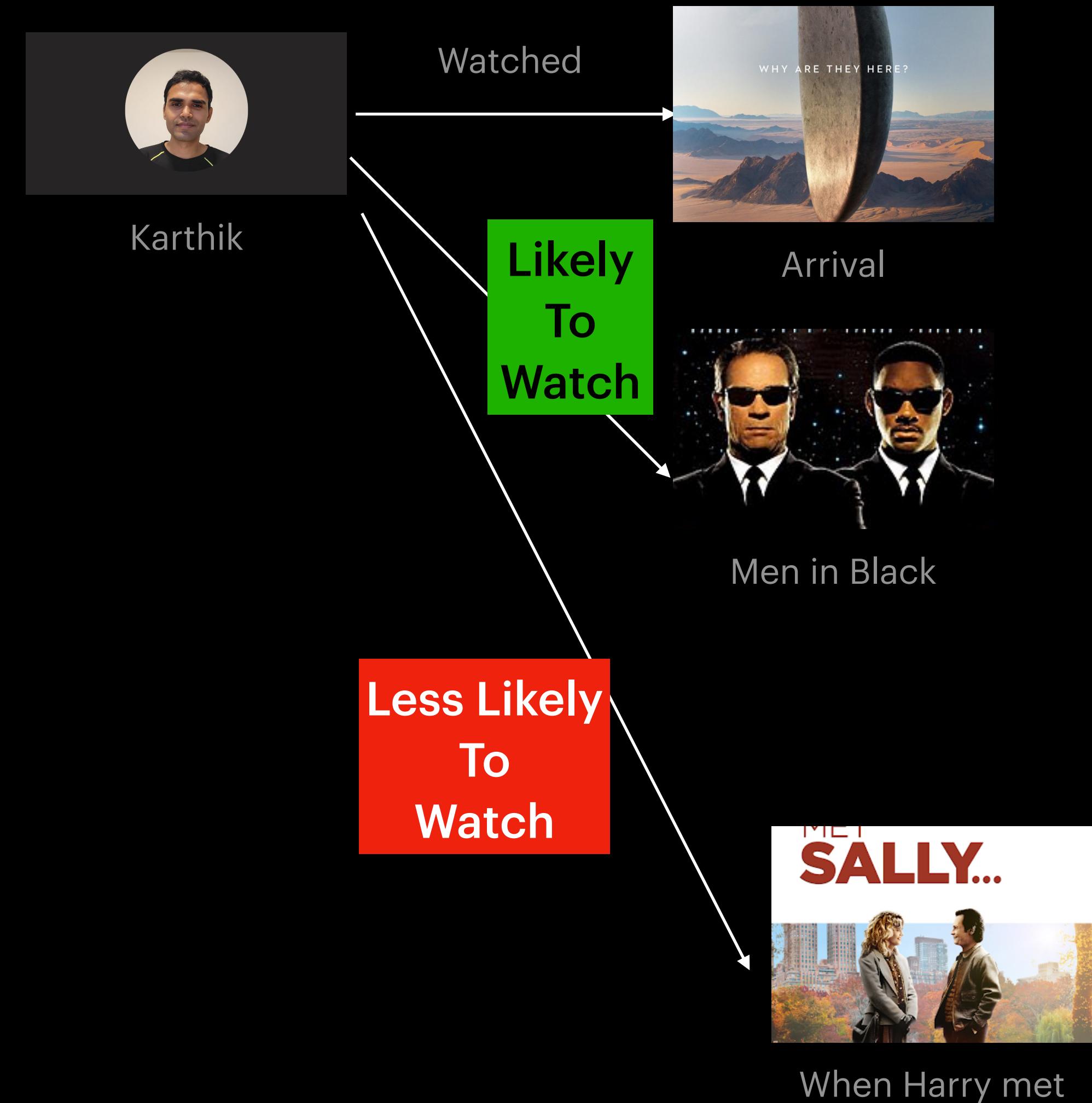
# Cauchy-Schwarz Inequality!

$$-1 \leq \text{CosineSimilarity}(x, y) = \frac{x^T y}{\|x\| \|y\|} \leq 1$$

$$|x^T y| \leq \|x\| \|y\|$$

$\|x\|_2$  Euclidean Norm of x

# Embeddings in Action



**What if I like both sci-fi and romance?**

# Content Based Filtering



Minions



Men in Black



Men in Black



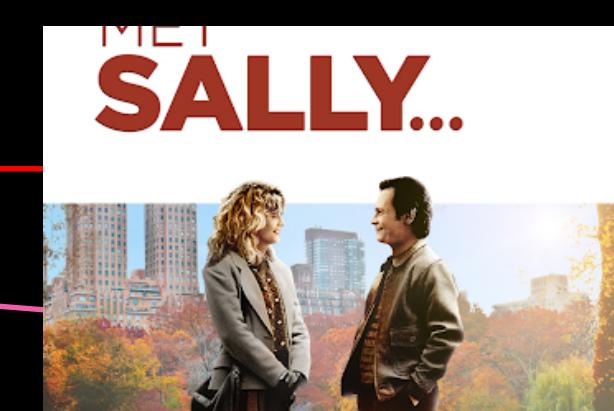
Arrival



When Harry met

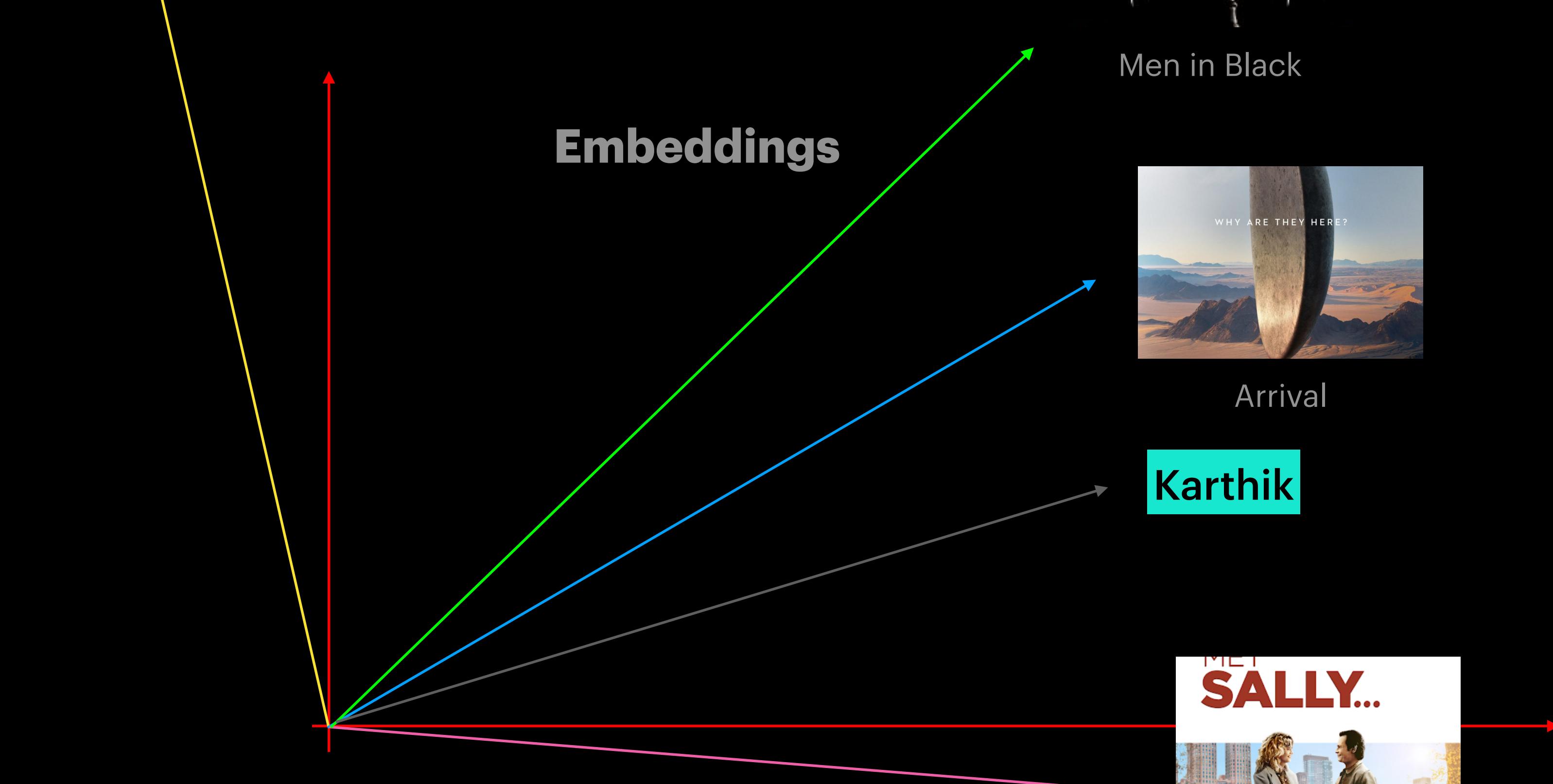


Karthik

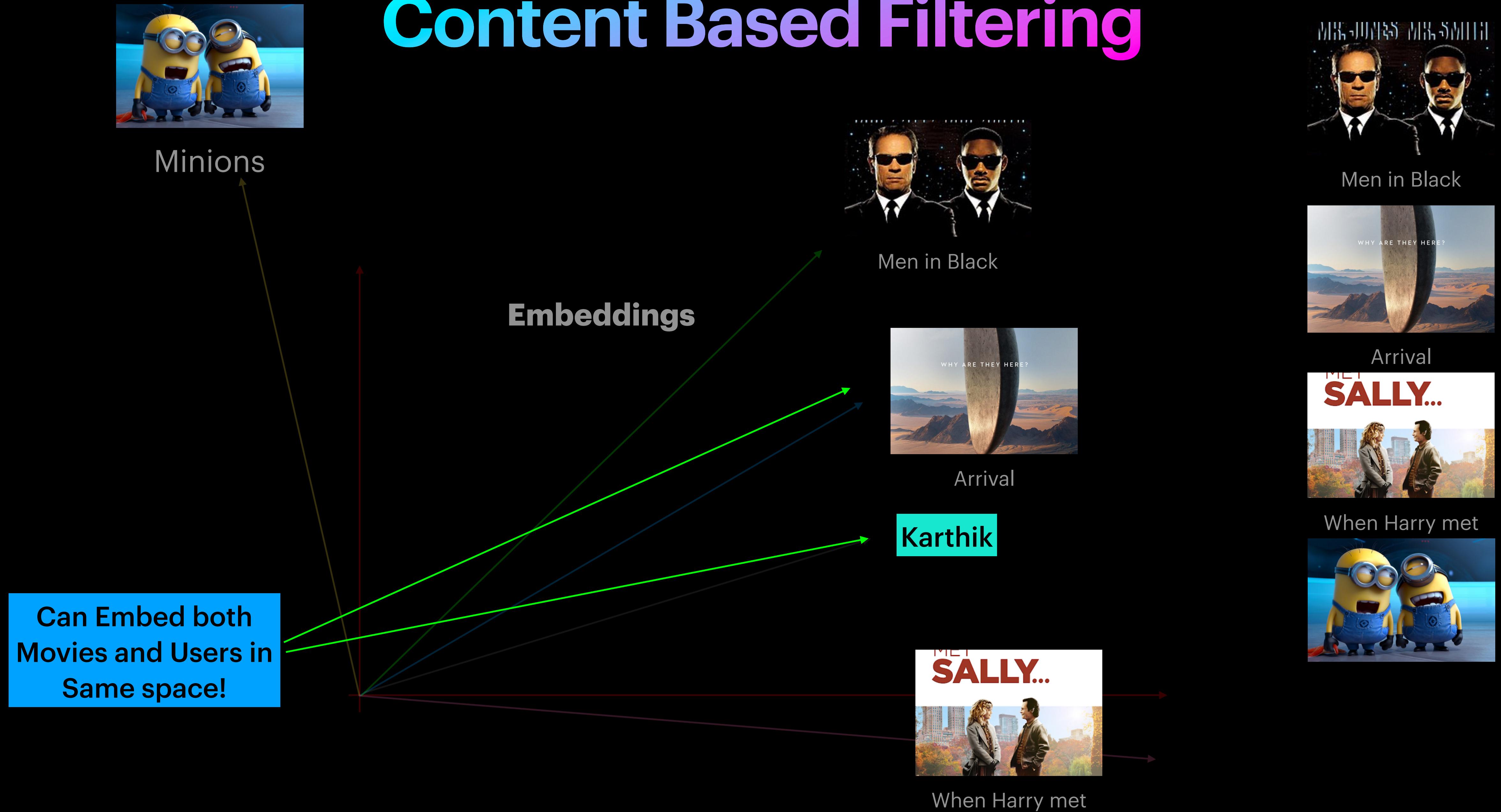


When Harry met

Embeddings



# Content Based Filtering



# Content Based Filtering



Minions



Men in Black



Men in Black



Arrival

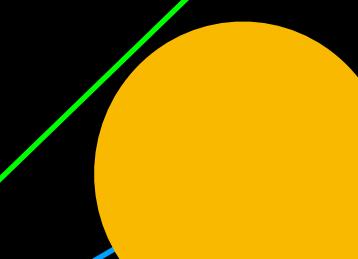


When Harry met

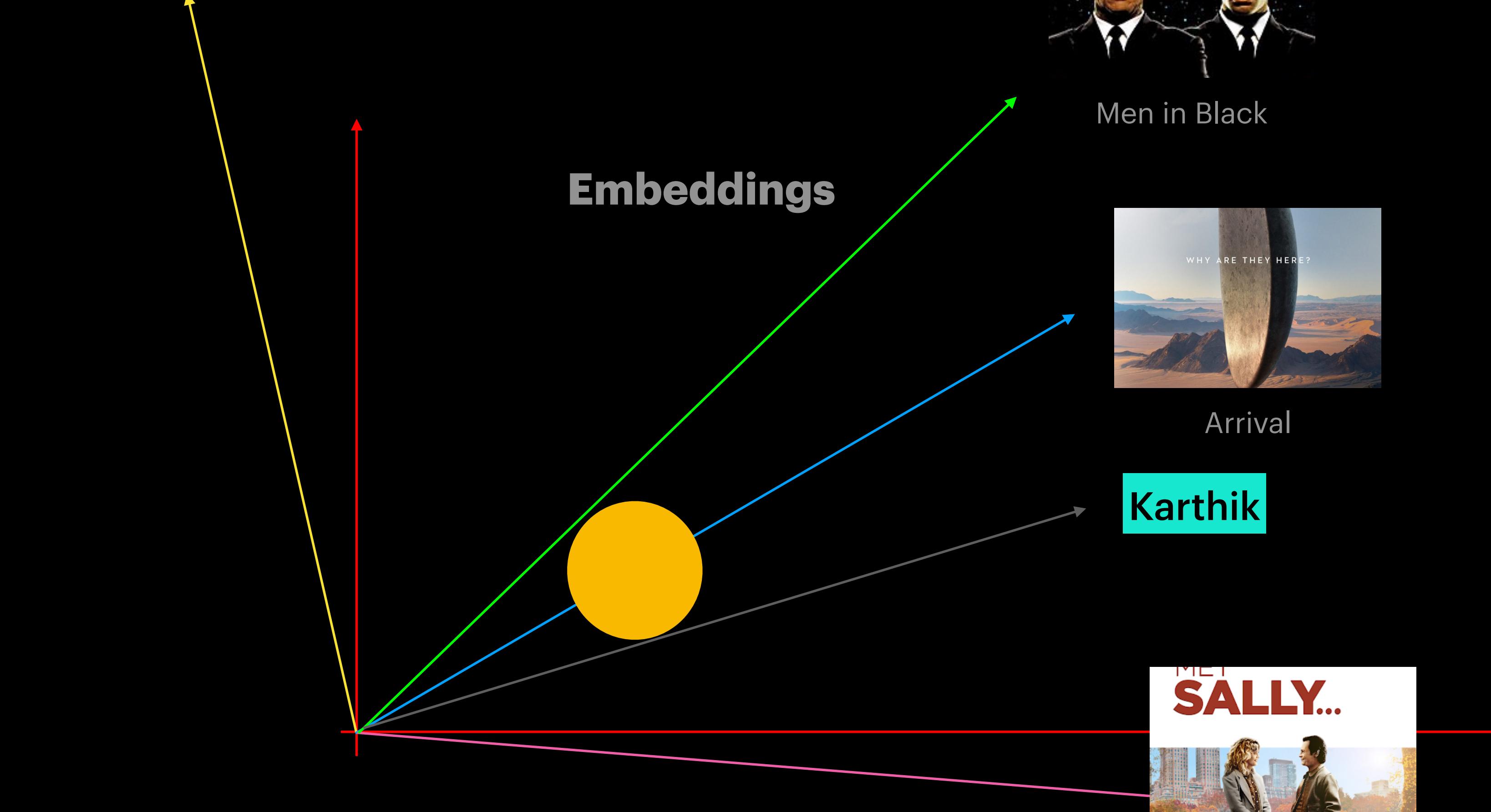


When Harry met

Embeddings



Karthik



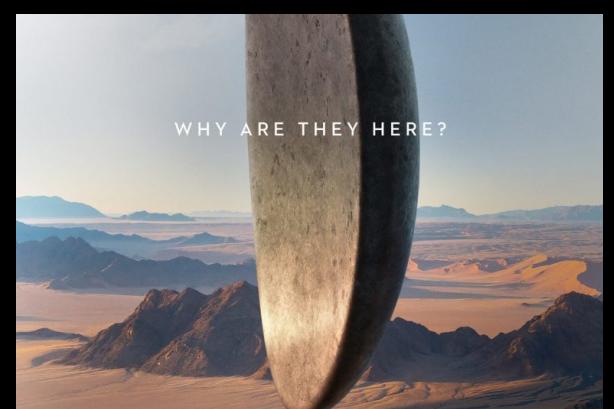
# Content Based Filtering



Minions



Men in Black



Arrival



Arrival



When Harry met

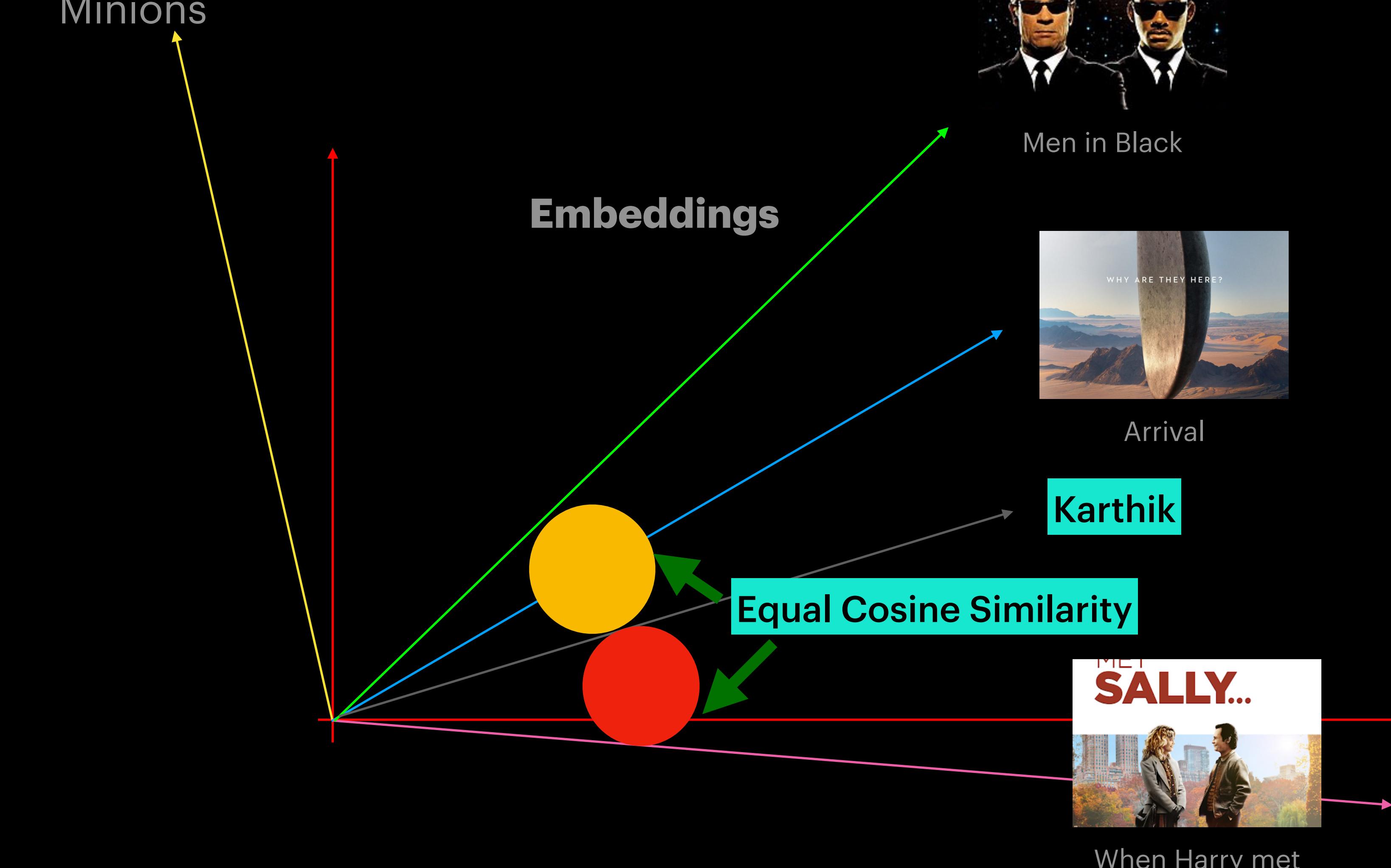


When Harry met

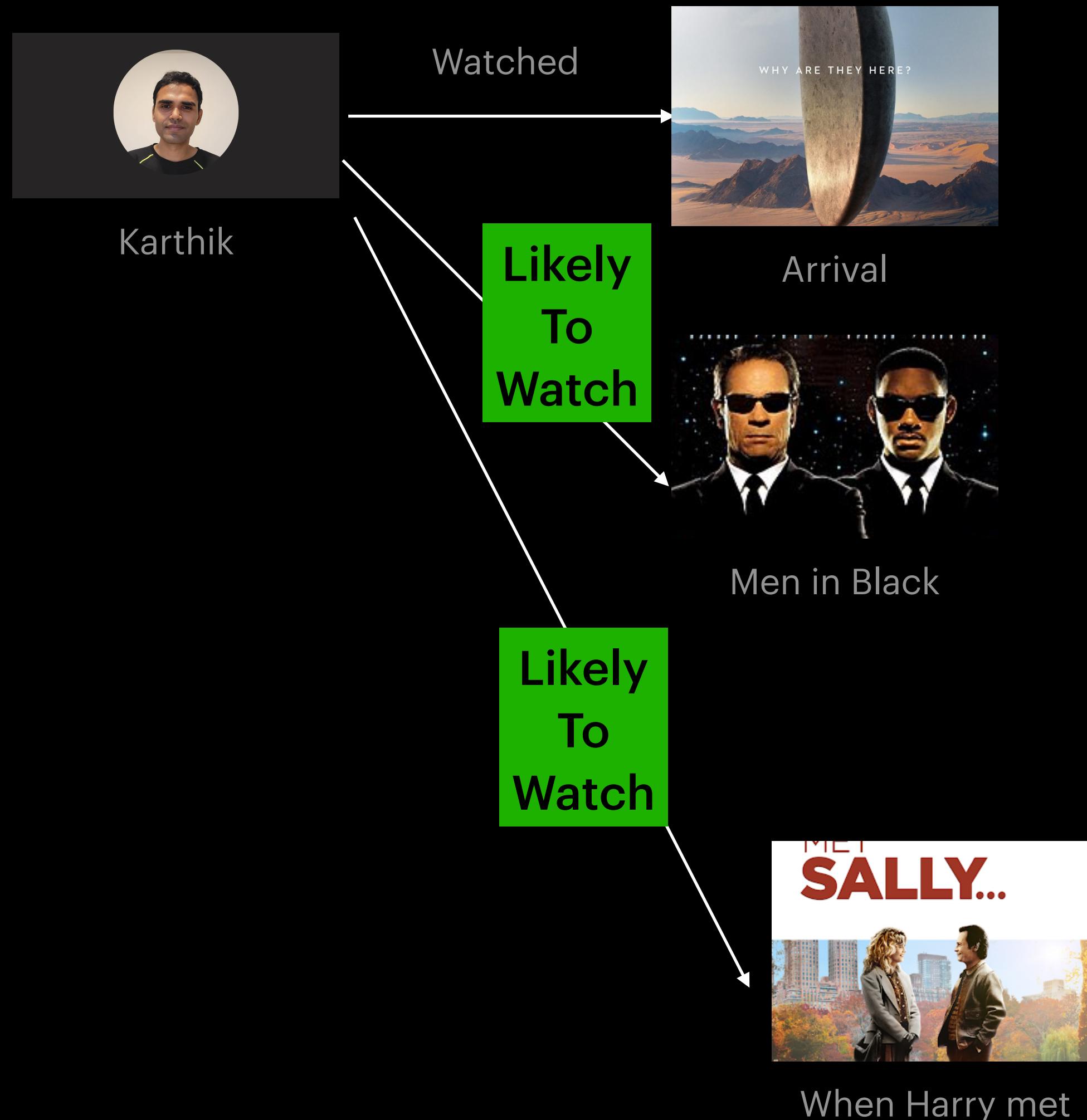
Embeddings

Karthik

Equal Cosine Similarity



# Like both Sci-fi and Romance



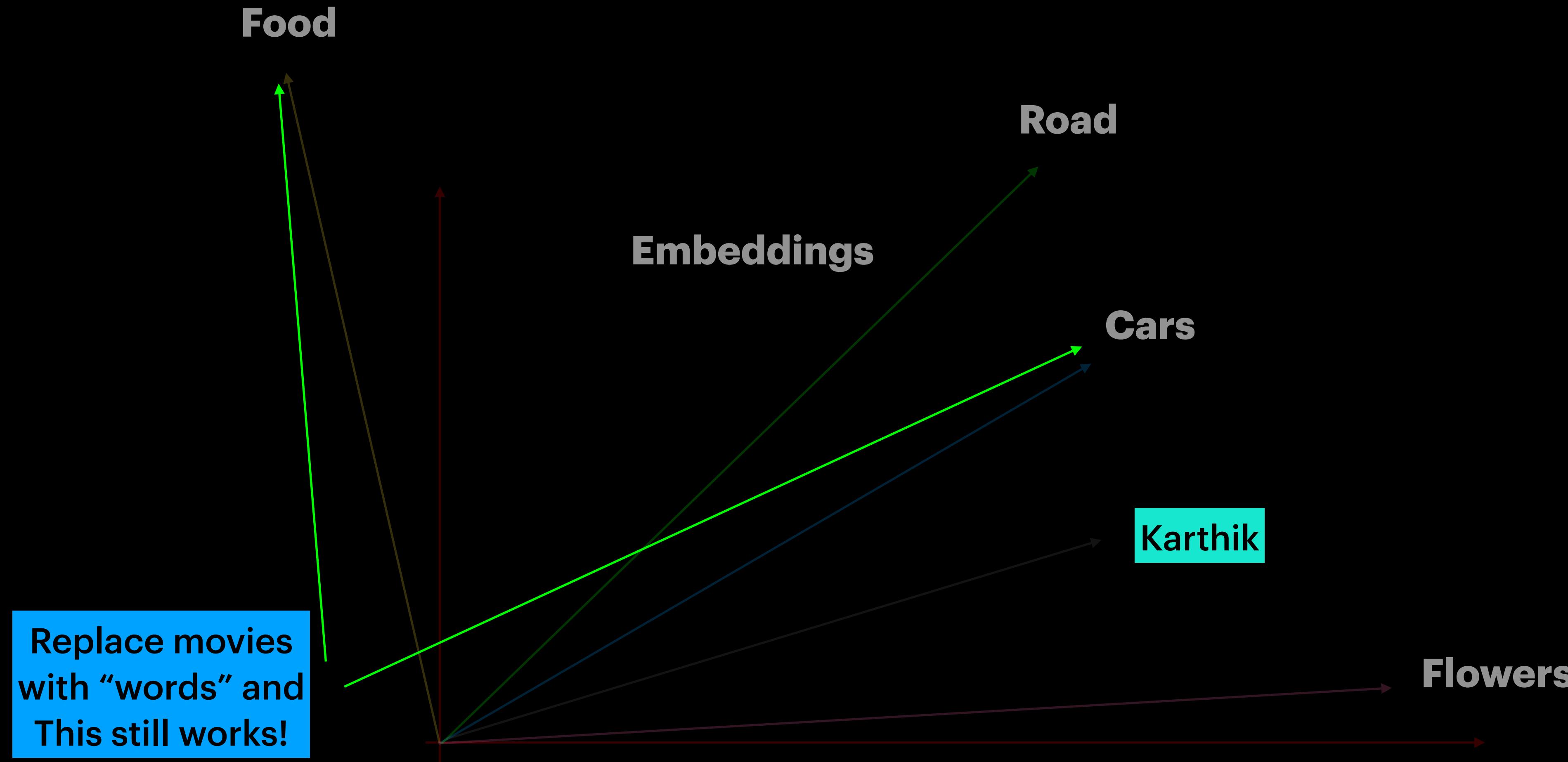
# ICE #4

$$\text{CosineSimilarity}(x, y) = \frac{x^T y}{\|x\| \|y\|}$$

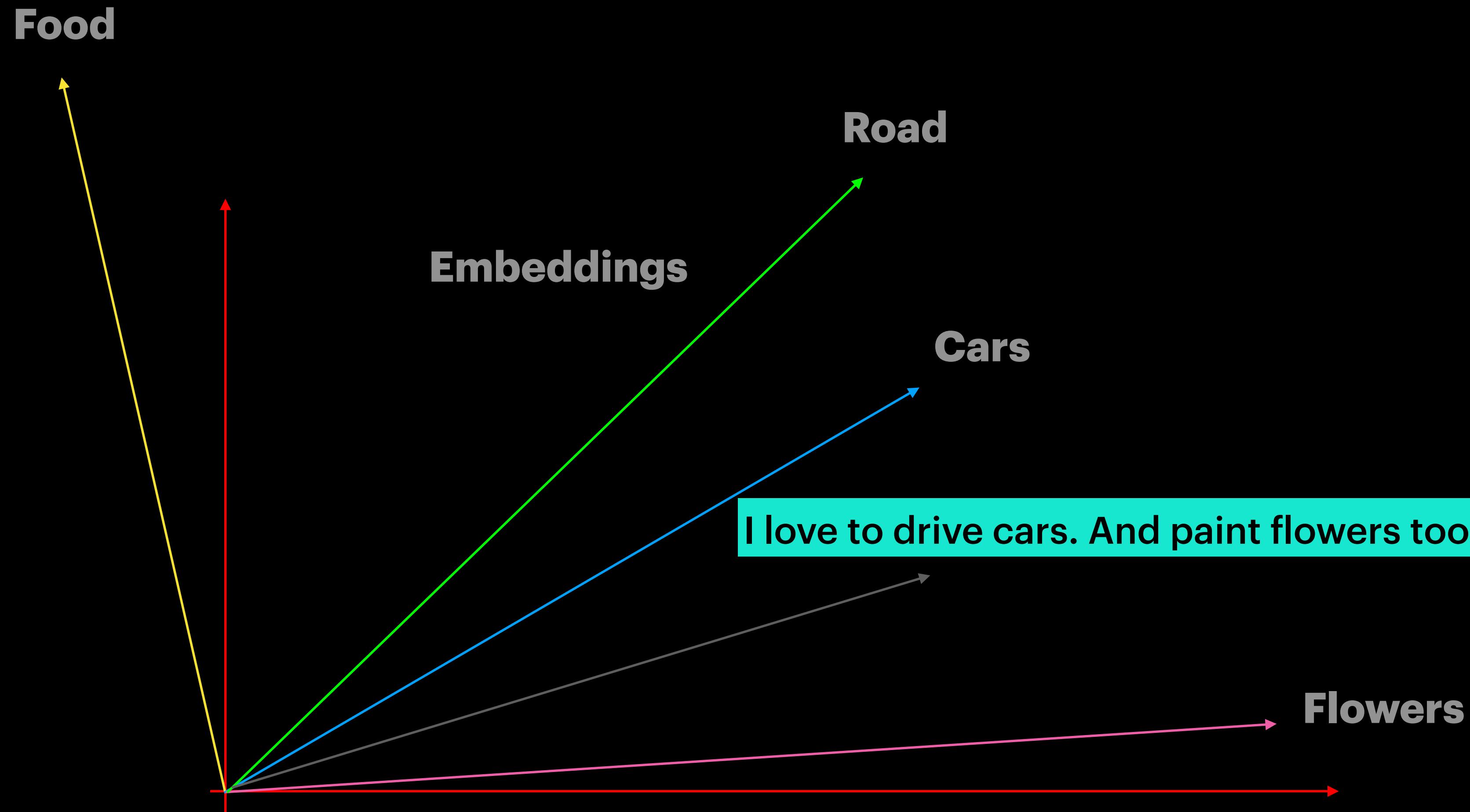
Lets say that **Avatar** had a 4d embedding as **[1,-1,2,5]** and **Michael** had an embedding of **[-1,2,1,-1]**. Should we recommend **Avatar** to **Michael**?

- a) Yes
- b) No

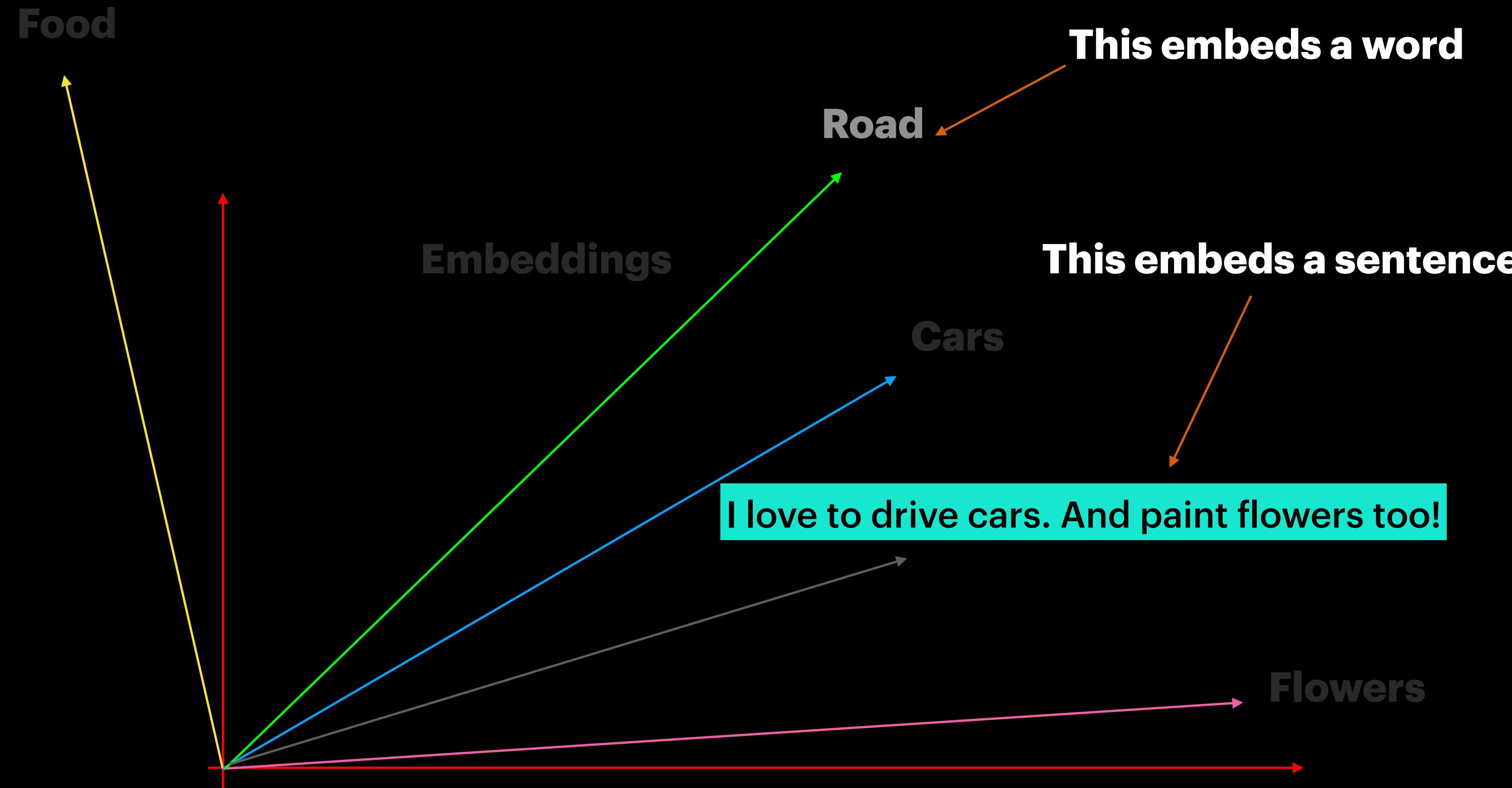
# Word Embeddings



# Word Embeddings



# Word and Sentence Embeddings



# Word and Sentence Embeddings

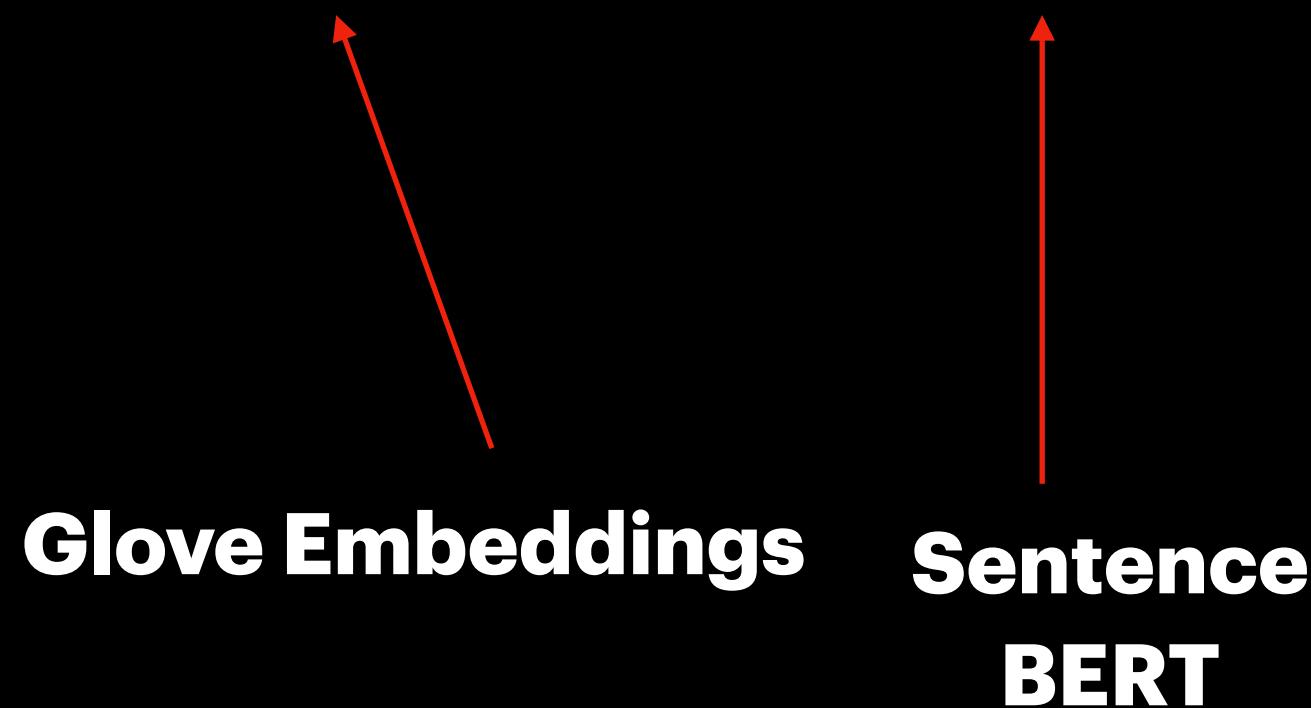
**How do we  
Obtain these embeddings?**

**A: Through a DL model!  
Maybe last but one hidden layer activations**

# Word and Sentence Embeddings

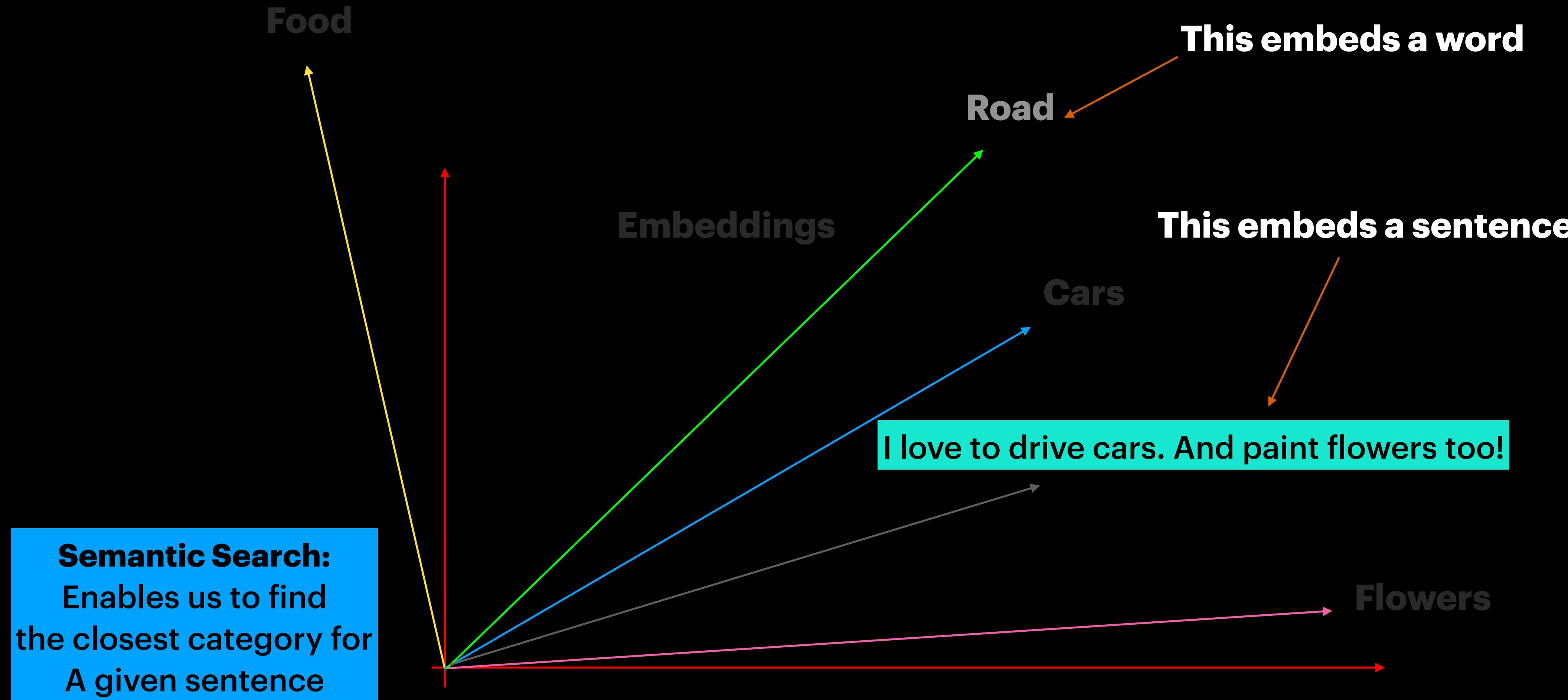
How do we  
Obtain these embeddings?

A: Through a DL model!  
Maybe last but one hidden layer activations

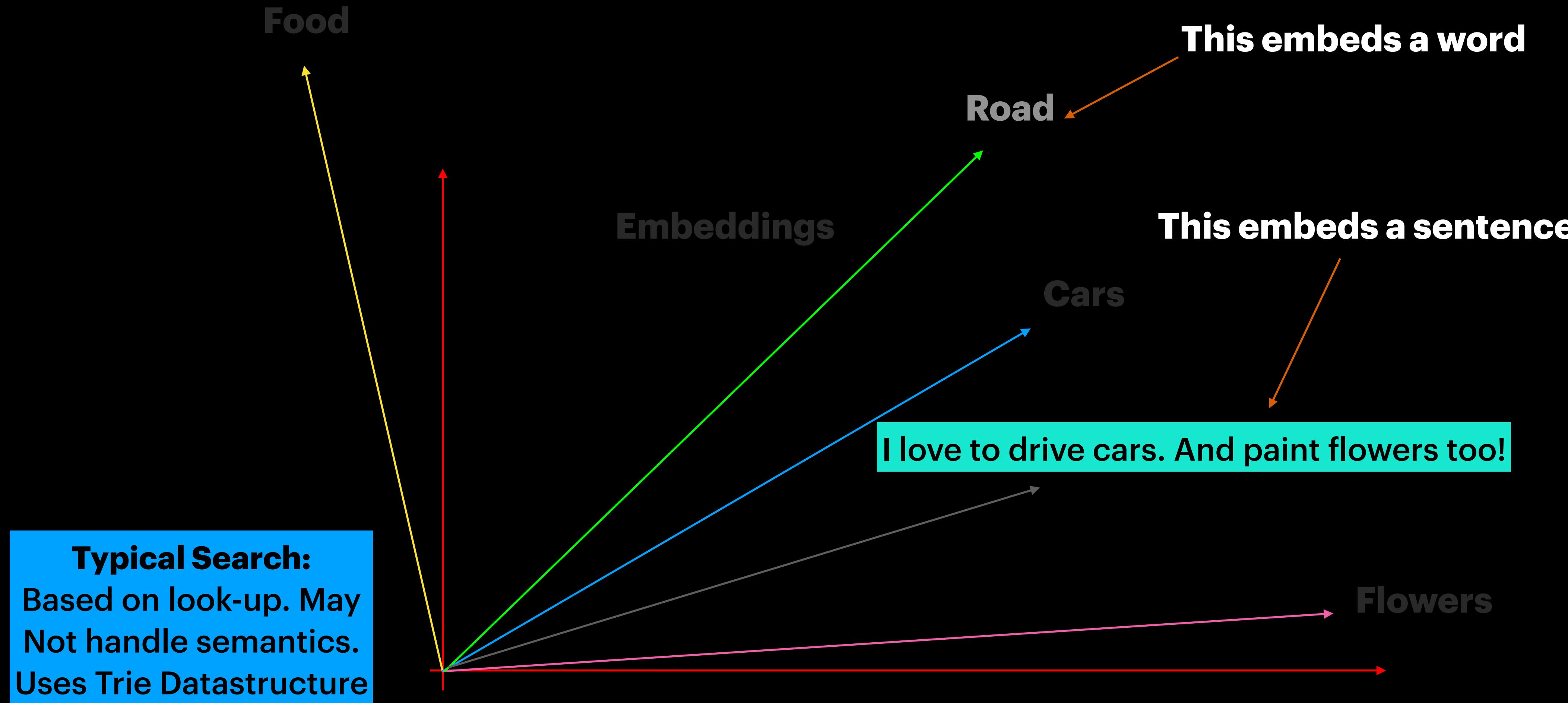


We will cover Sentence BERT when we  
Get to Transformers!

# Semantic Search | Vector Search



# Semantic Search | Vector Search



# Semantic Search | Vector Search

# Food

# Embeddings

# Road

# This embeds a word

# This embeds a sentence

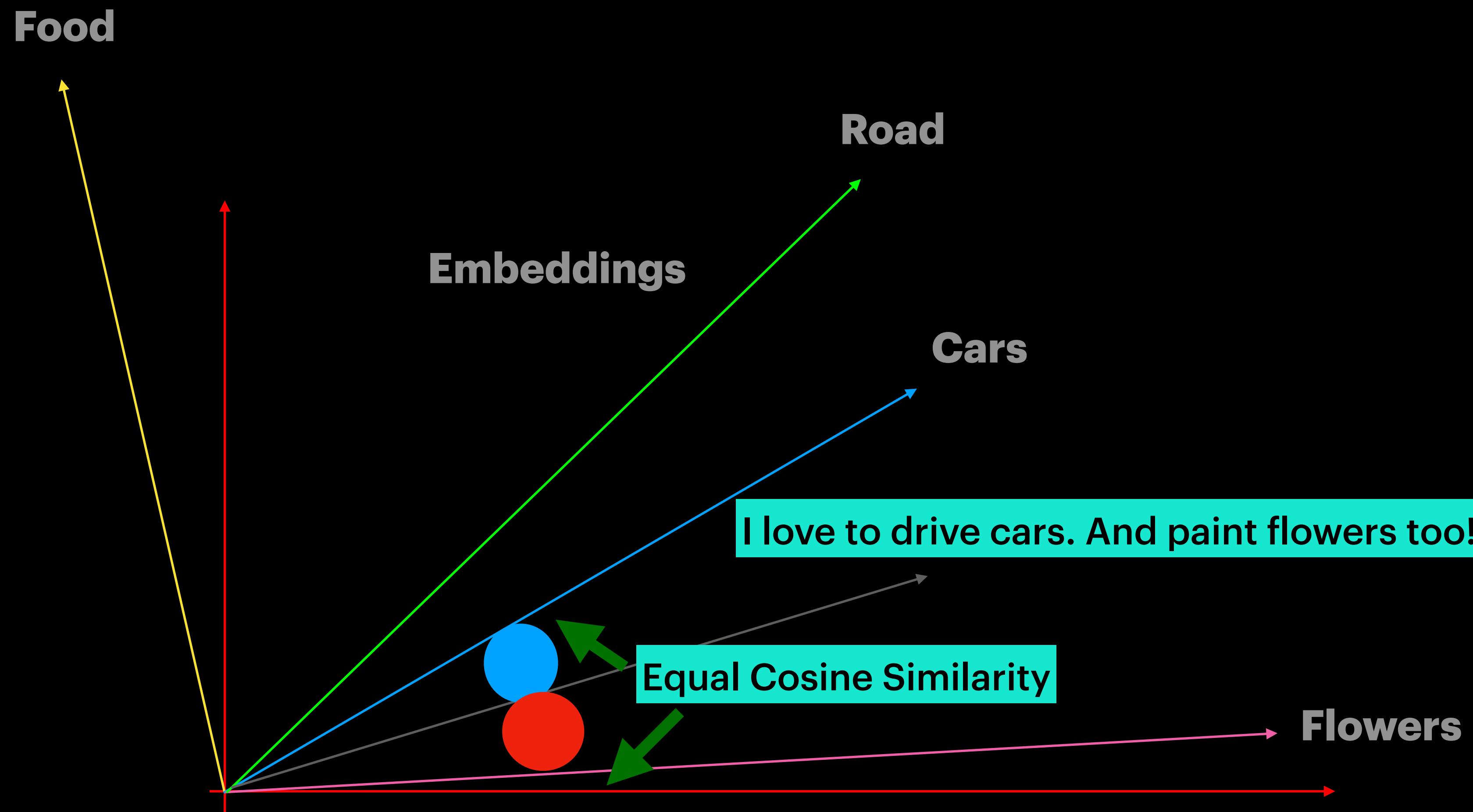
# Cars

I love to drive cars. And paint flowers too!

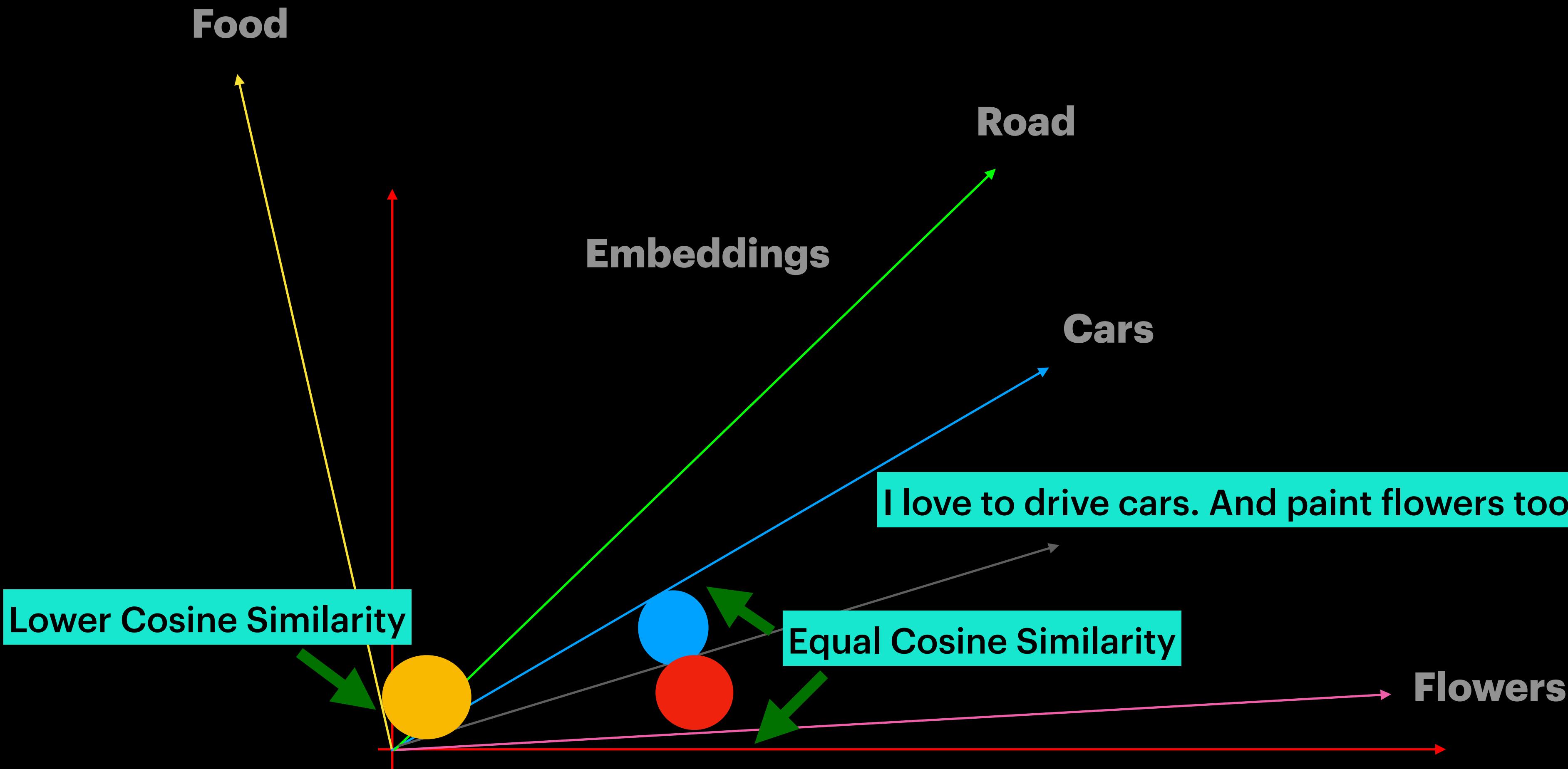
# Typical Search: Based on look-up.

# Down-sides?

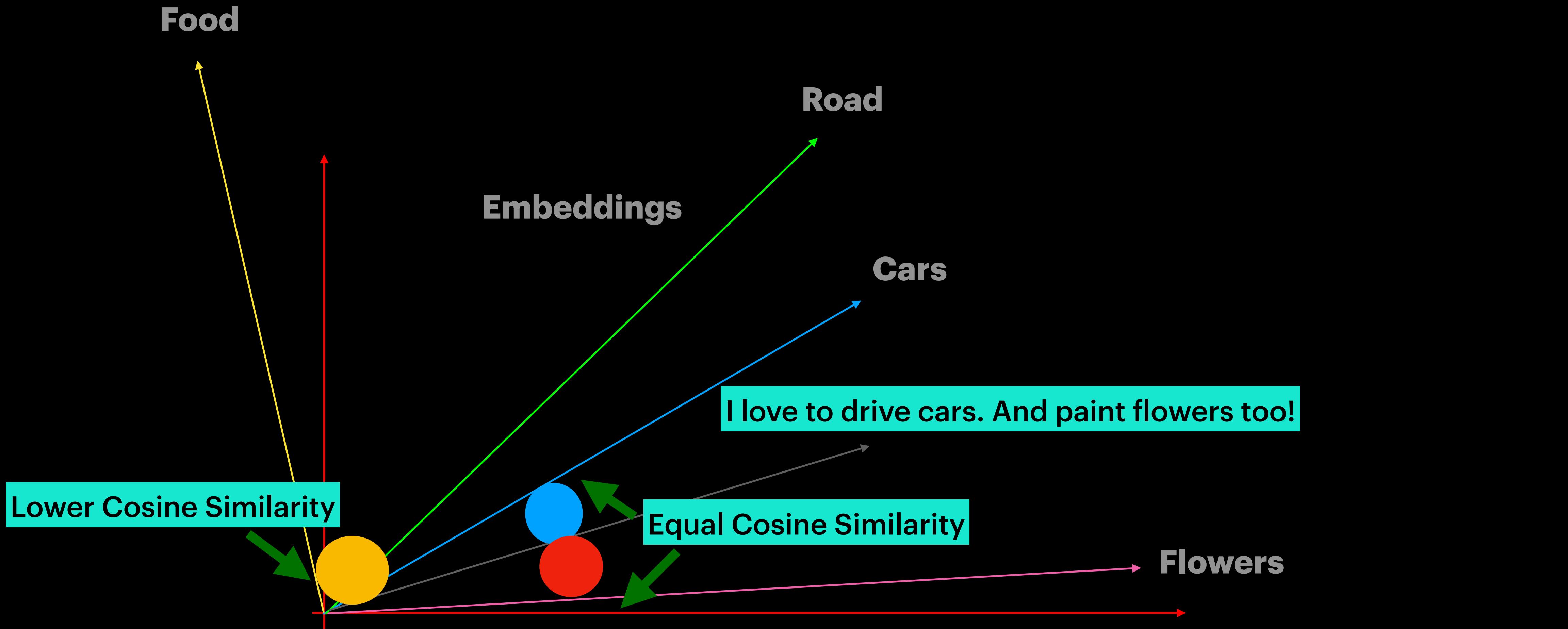
# Semantic Search | Vector Search



# Semantic Search | Vector Search



# Semantic Search | Vector Search



What's the closest category  
for the following sentence? "I love to drive cars. And paint flowers too!"

Highest cosine similarity based on  
vector search: Flowers and Cars

# Vector Arithmetic!

What is King - Man + Woman?

# Demo on Semantic Search