

# Making recommendations explainable

Lev Fedorov

# Introduction

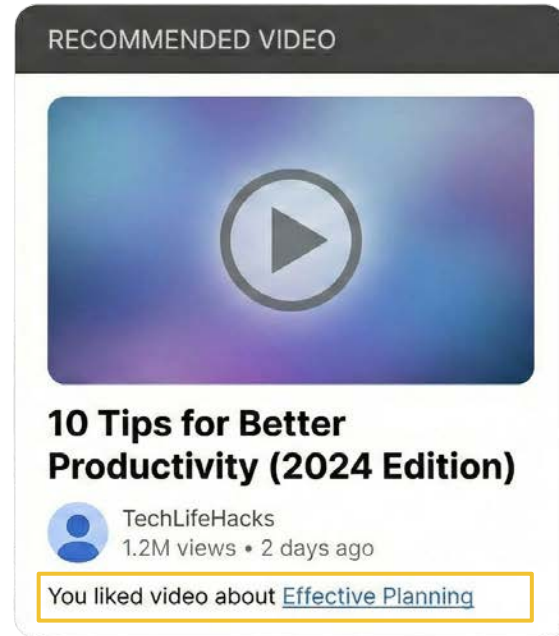
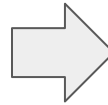
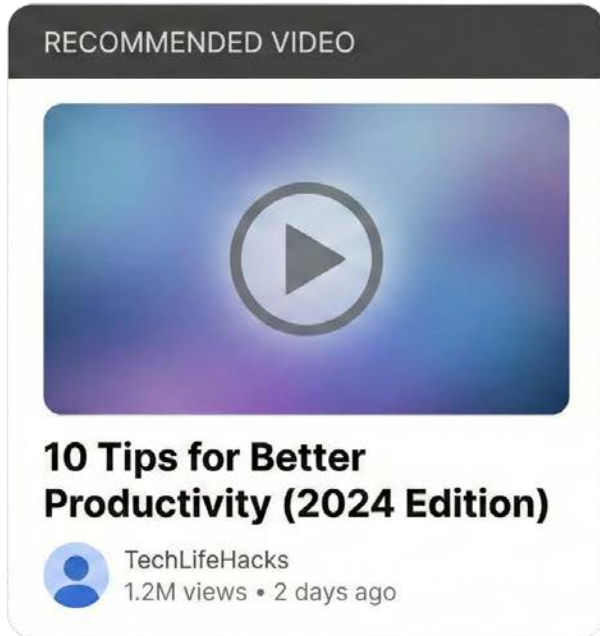
# Recommendation system



# Recommendation feed

The image shows a YouTube recommendation feed on a dark background. At the top, there is a navigation bar with the YouTube logo, a search bar, and a 'Sign in' button. Below this is a horizontal menu of category buttons: 'All', 'Dog grooming', 'Podcasts', 'Retrievers', 'Gaming', 'Conformation Shows', 'AI', 'Seasons', 'News', 'Dog barks', 'Live', 'Meow', and 'Budget'. The main content area displays a grid of video thumbnails. The first row includes 'Tom and Jerry | Mega Episode: Golden Era Vol. 1 | Warner Classics' (55:48), 'Top Hits 2025 Playlist ~ Trending Music 2025 Spotify Mix ~ Best TikTok Songs ...' (2:32:57), and 'Best of 2022 Masters Agility Championships from Westminster Kennel...' (13:31). The second row includes '5 Times VERY Persistent Dogs Convinced' (13:56), 'Jürgen Klopp x AlphaTauri' (16:52), and 'The Job Market Has Changed' (16:52). On the left side, there is a sidebar with navigation options: 'Home', 'Shorts', 'You', 'History', 'Sign in to like videos, comment and subscribe.', 'Sign in', 'Explore', 'Music', 'Movies & TV', 'Live', 'Show more', 'More from YouTube', 'YouTube Premium', and 'YouTube Music'.

# Explanation



# Motivation

# Problems

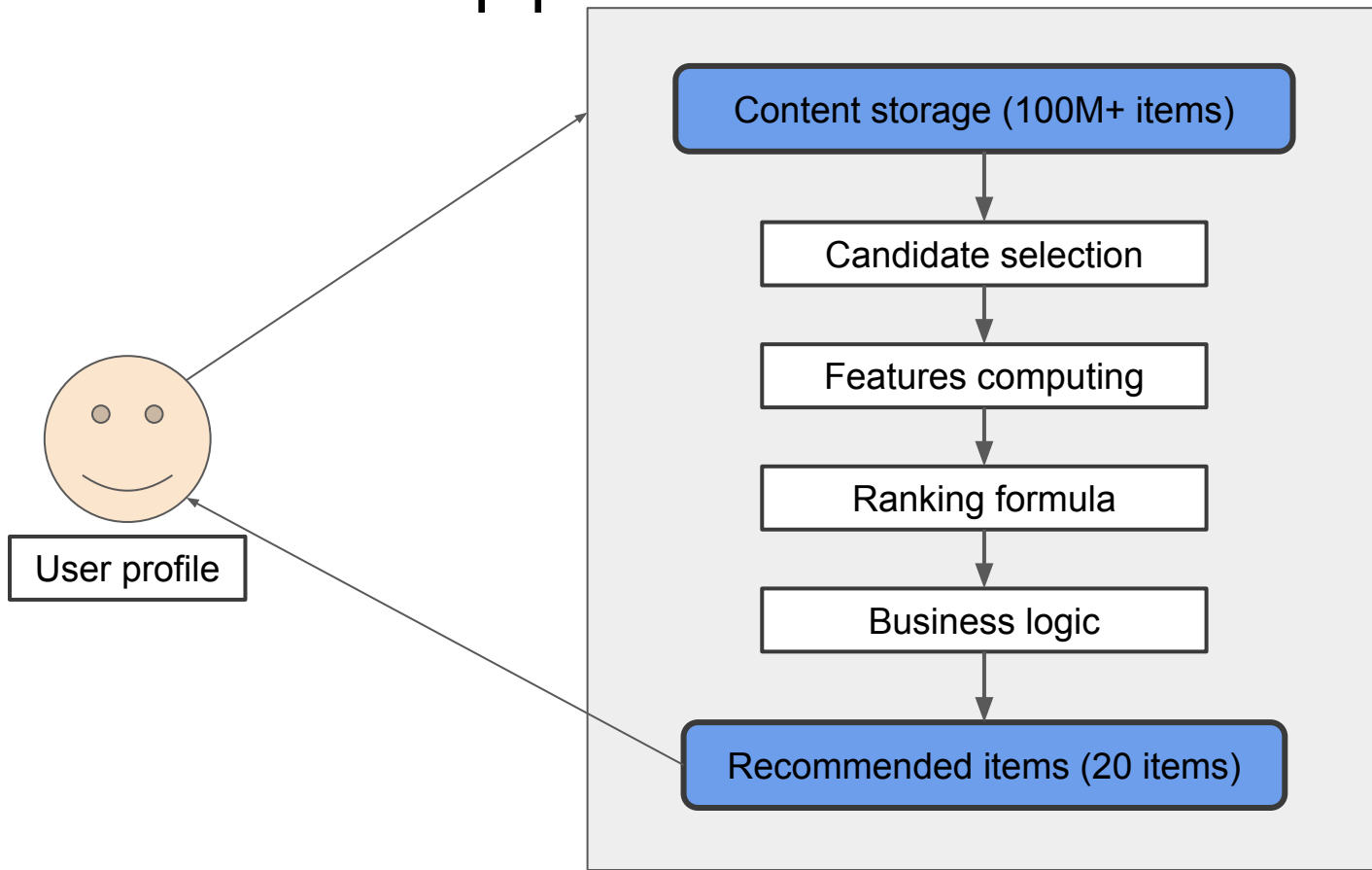
## Users

- Reports about feed irrelevancy

## Developers

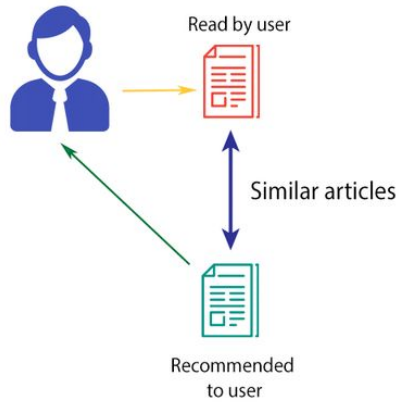
- No traceability/debugging tools

# Recommendations pipeline

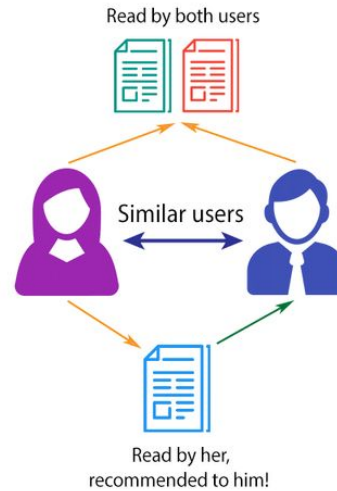


# Recommendation systems

CONTENT-BASED FILTERING



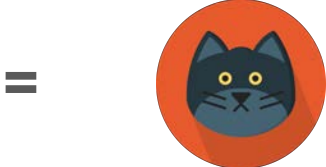
COLLABORATIVE FILTERING



# Embedding

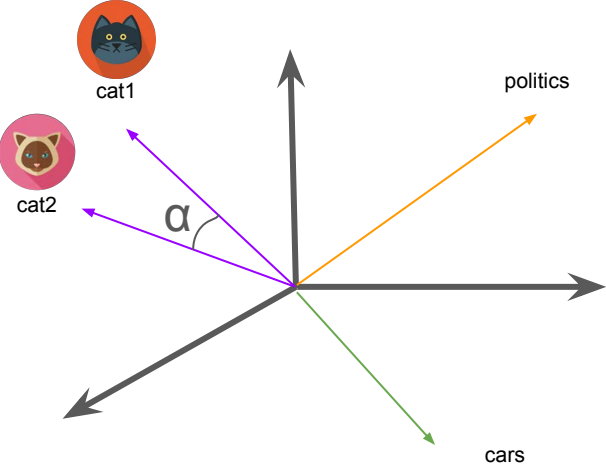
embedding = float[] array = vector

1	2	...	N - 1	N
0.98	0.22	...	0.45	0.78



$\angle \alpha \rightarrow 0^\circ$

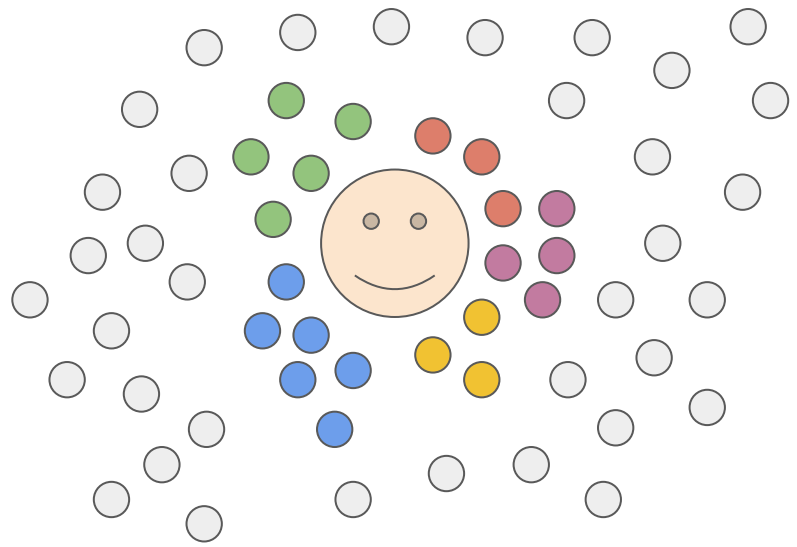
$\cos(\alpha) \rightarrow 1$



expectations

# User profile / One embedding

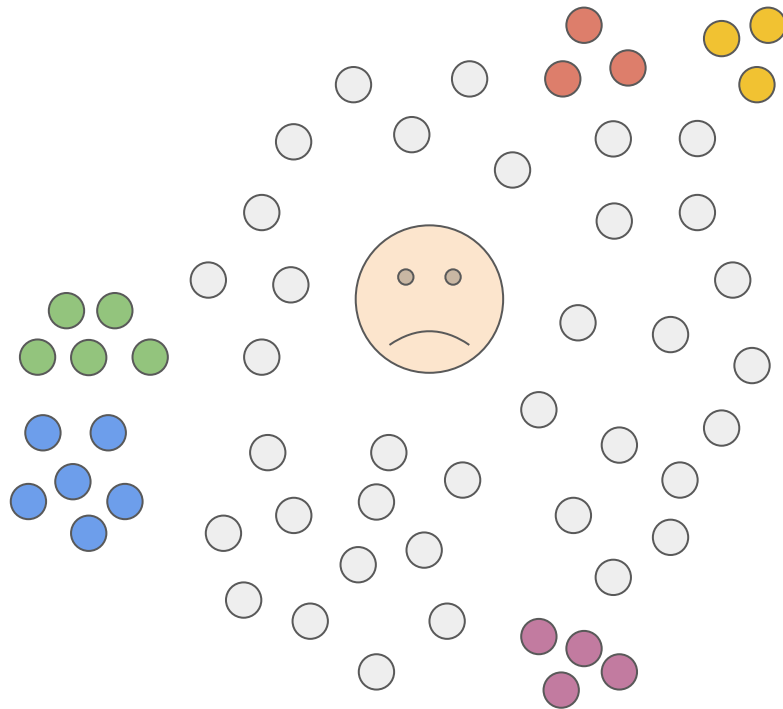
- - cats
- - dogs
- - cars
- - it
- - fashion
- - other
- 😊 - user



# User profile / One embedding

reality

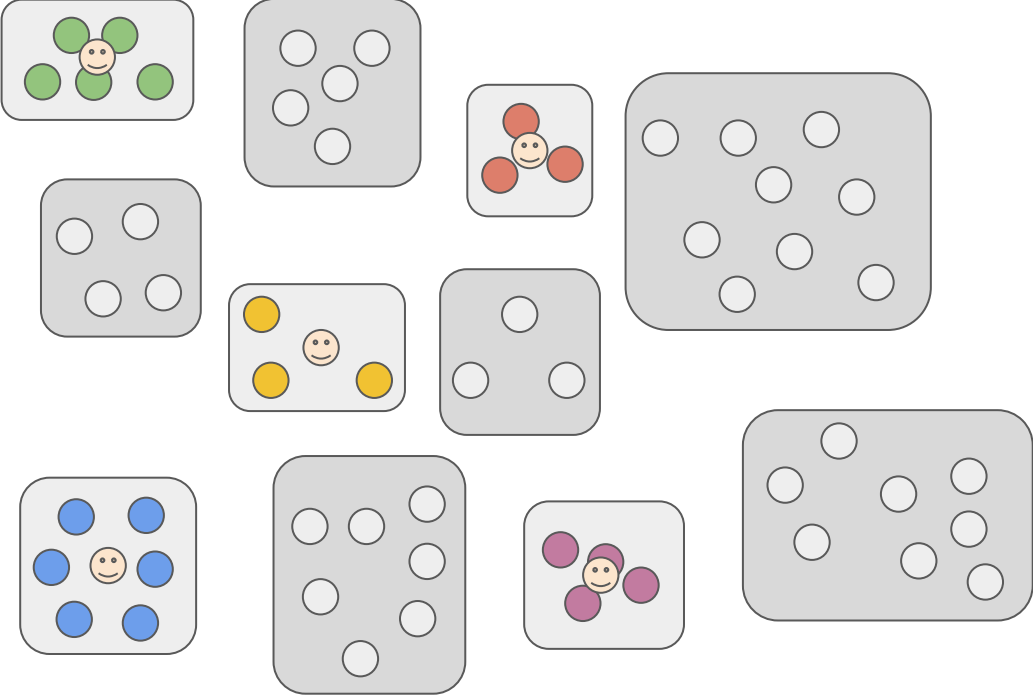
- - cats
- - dogs
- - cars
- - it
- - fashion
- - other
- ☹ - user



expectations

# User profile / K-Means (Centroids)

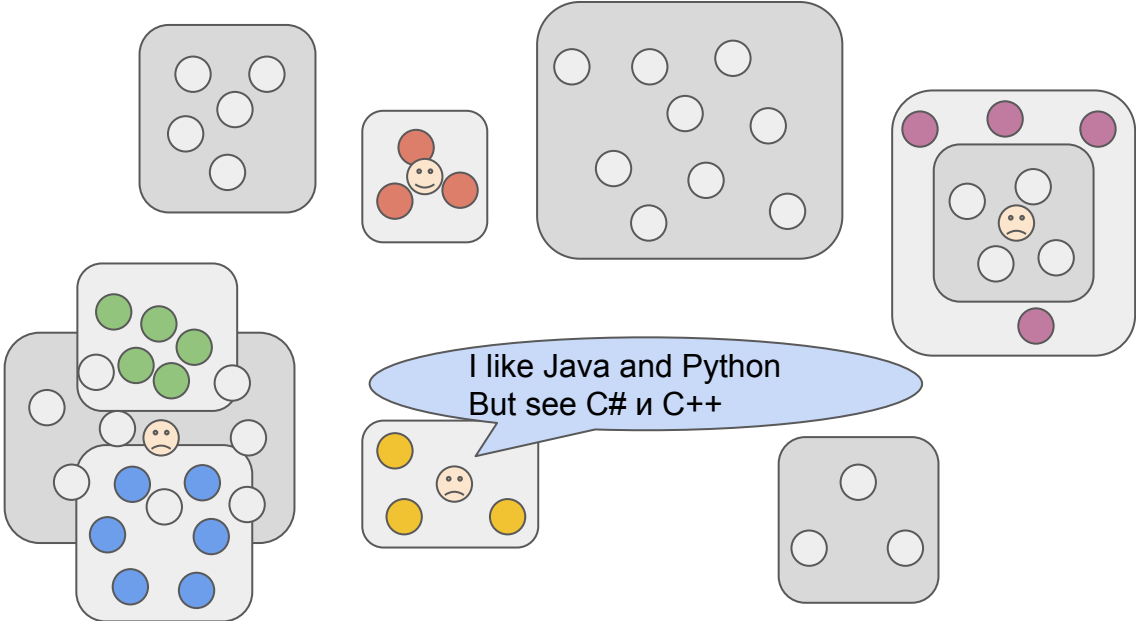
- - cats
- - dogs
- - cars
- - it
- - fashion
- - other
- 😊 - user



reality

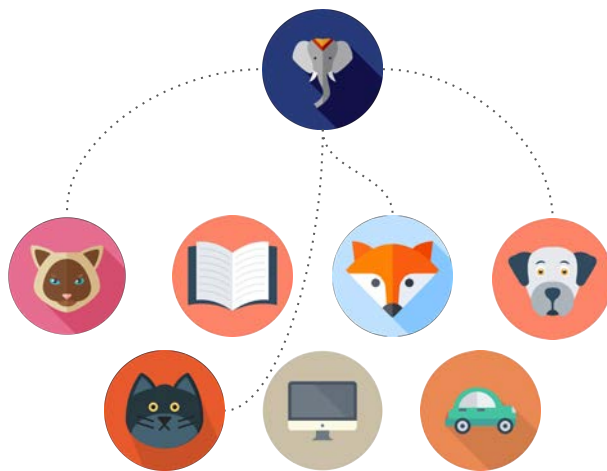
# User profile / K-Means (Centroids)

- - cats
- - dogs
- - cars
- - it
- - fashion
- - other
- ☹ - user

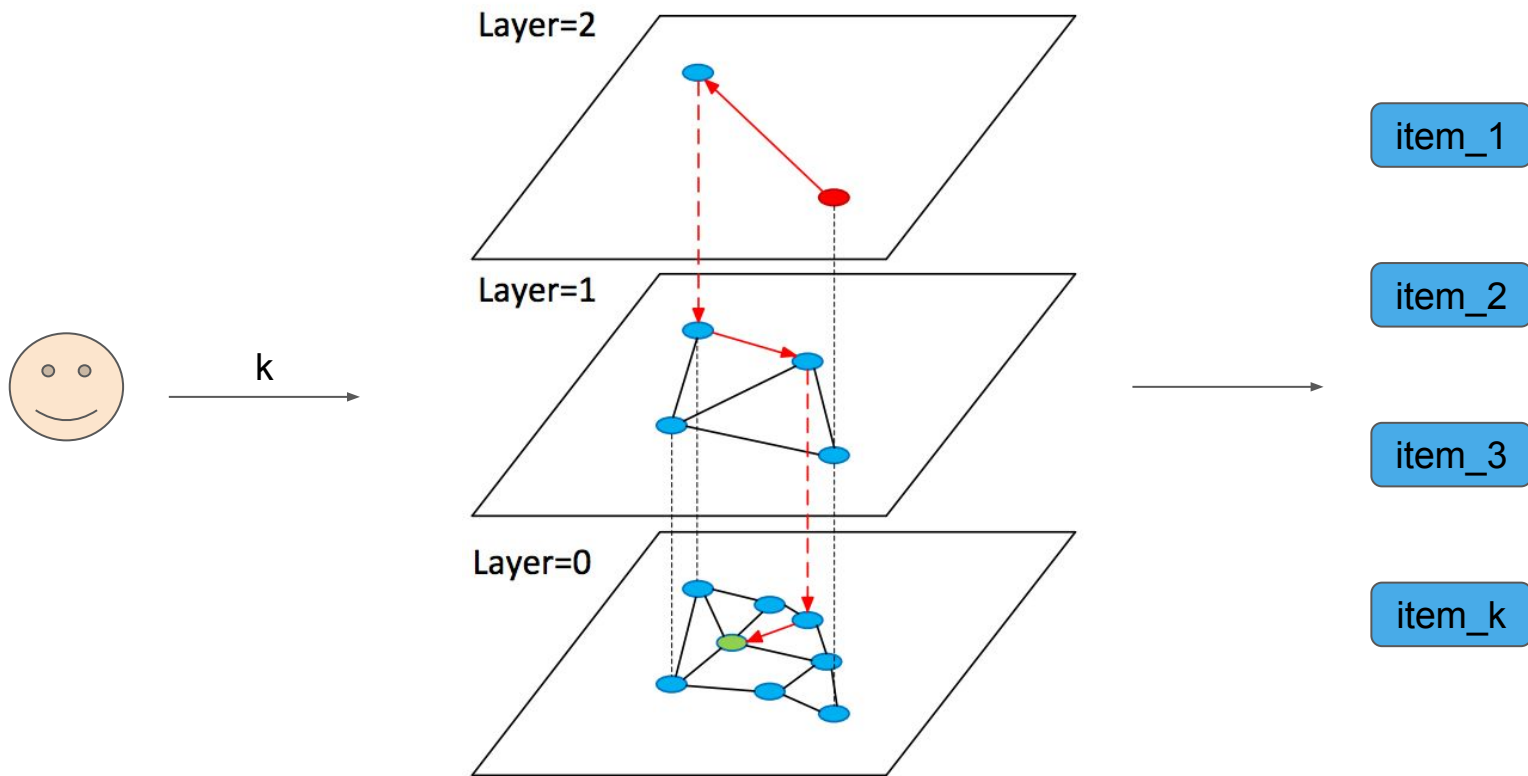


# Candidate selection / KNN

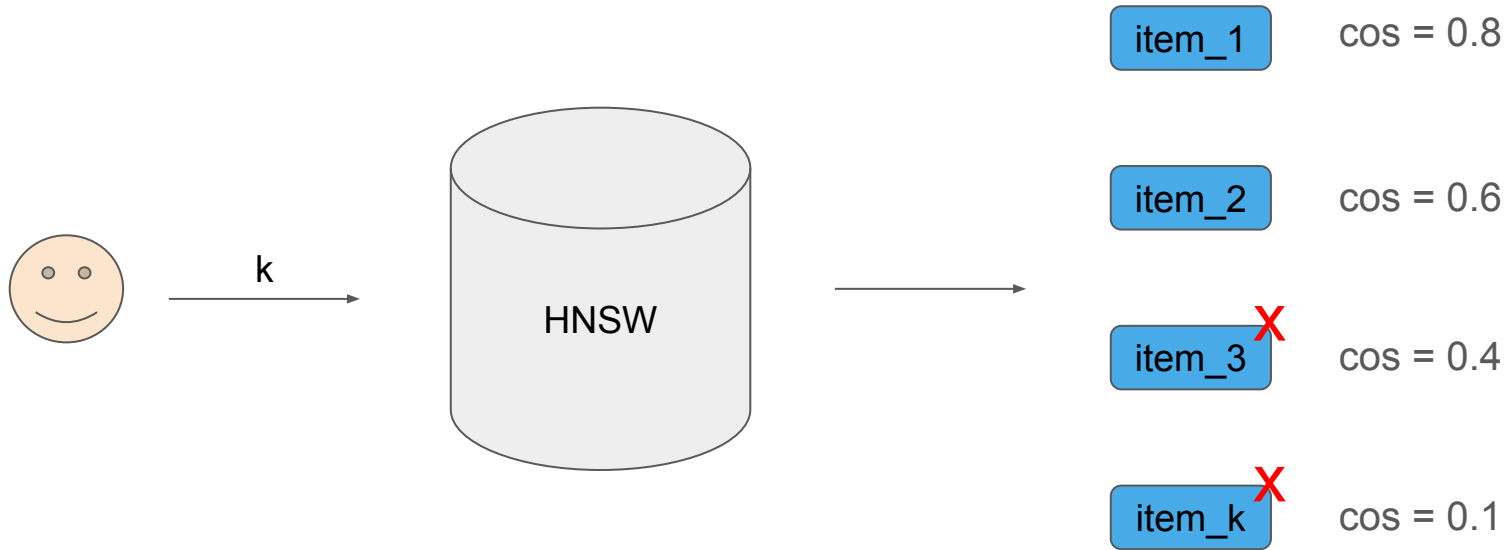
KNN (K-Nearest Neighbors) — return the K nearest neighbors to a given embedding



# KNN (HNSW)

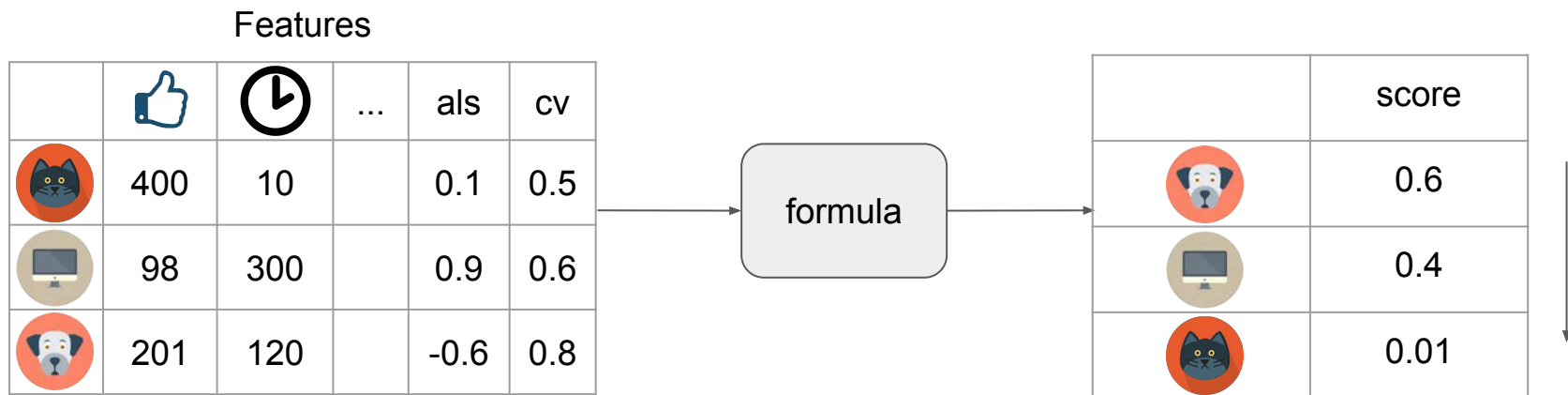


# KNN (HNSW)



# Ranking formula

Model (CatBoost) – predicts the publication relevance for user



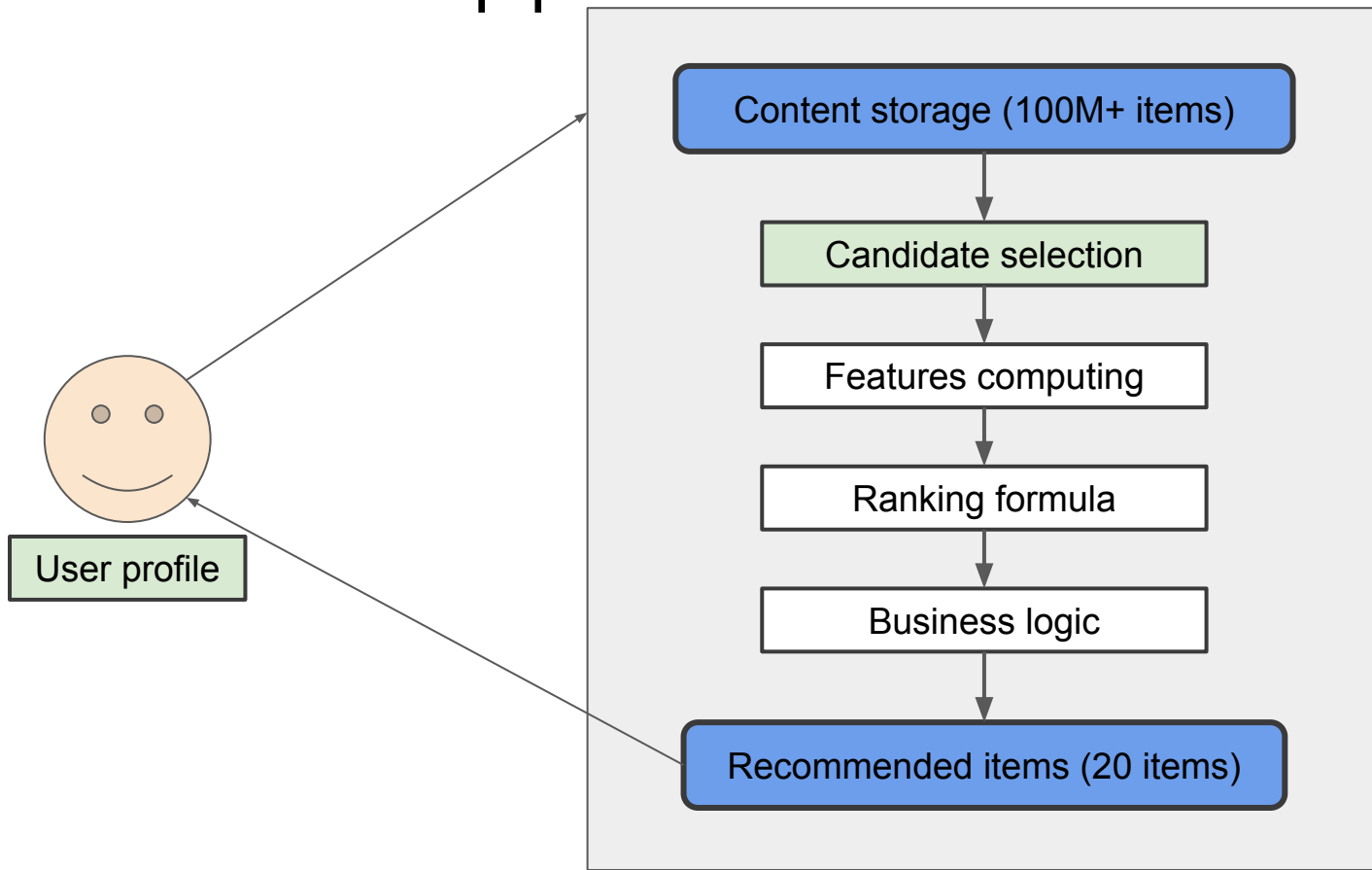
# Business logic

## Diversity

- Don't allow all posts to come from a single author
- Don't allow the entire feed to be only one topic/theme

# Approach

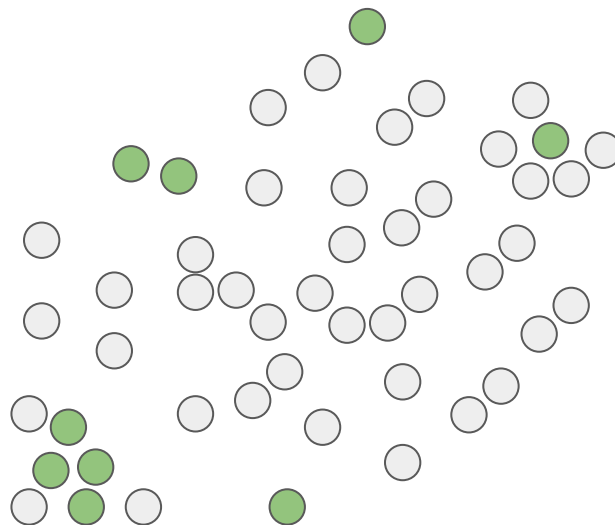
# Recommendations pipeline



# User profile

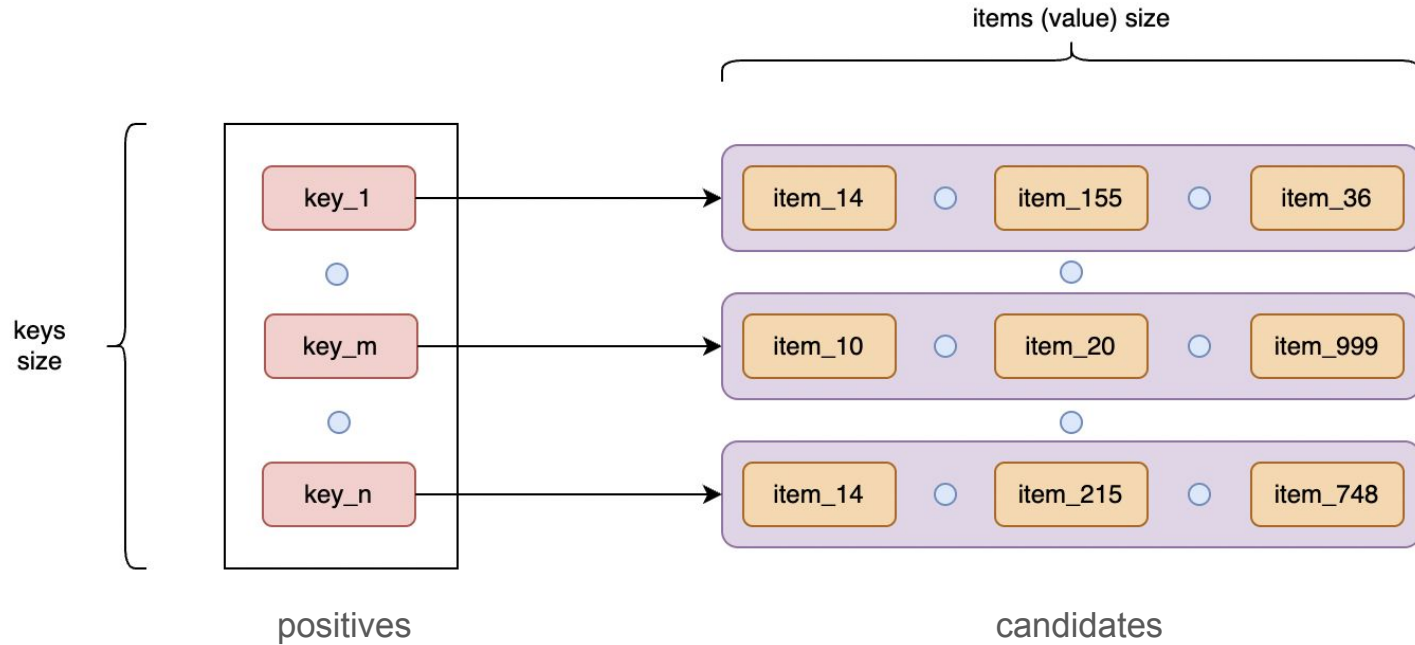
Instead of embedding, storing publication id

- - positive interaction
- - impression
- 😊 - user



😊 = [publication\_1, publication\_5, ...]

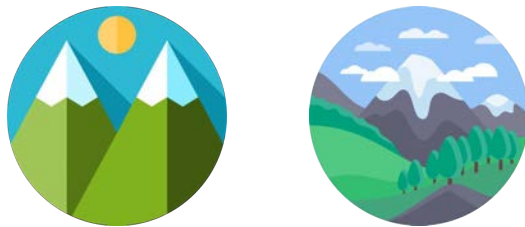
# Candidate Selection



# Relevance

Filter non-similar publications

- Annotate pairs of publications as “relevant” and “irrelevant”



are these publications similar?



- V0: Define cosine threshold for content-based model
- V1: Train relevance formula

# Attractiveness

Order list of candidates

- V1: Conditional CTR

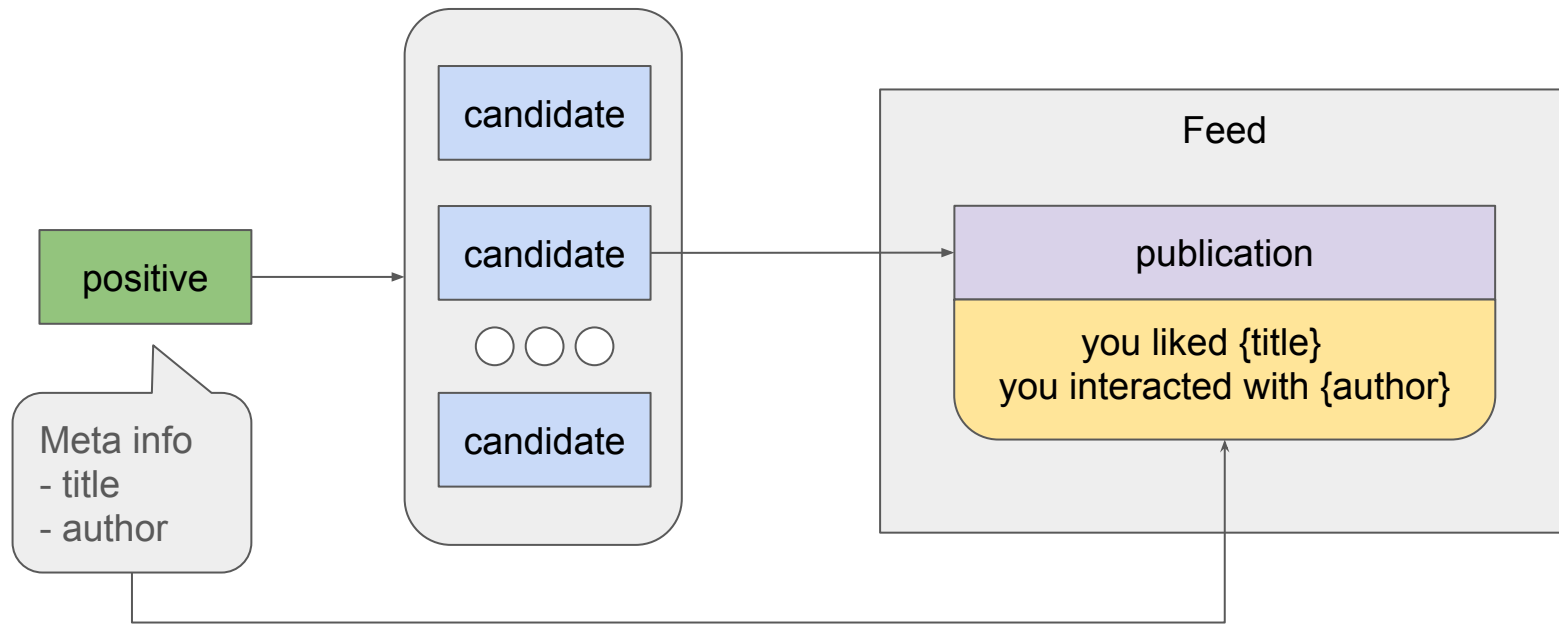
$$CTR(A|B) = CTR_A|click_B$$

Conditional CTR is the click-through rate of publication A given that the user had a positive interaction with publication B

- V2: Attractiveness formula

Predict positive interaction with publication A given that the user had a positive interaction with publication B

# Recommendations Pipeline



# Positives selection / DPP

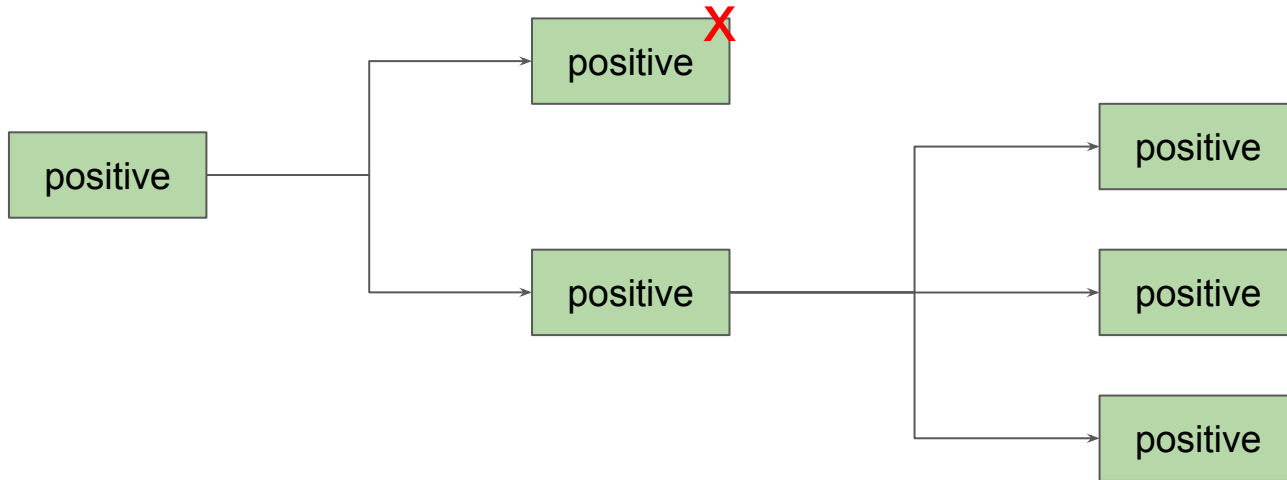
Determinantal Point Process

Select K points

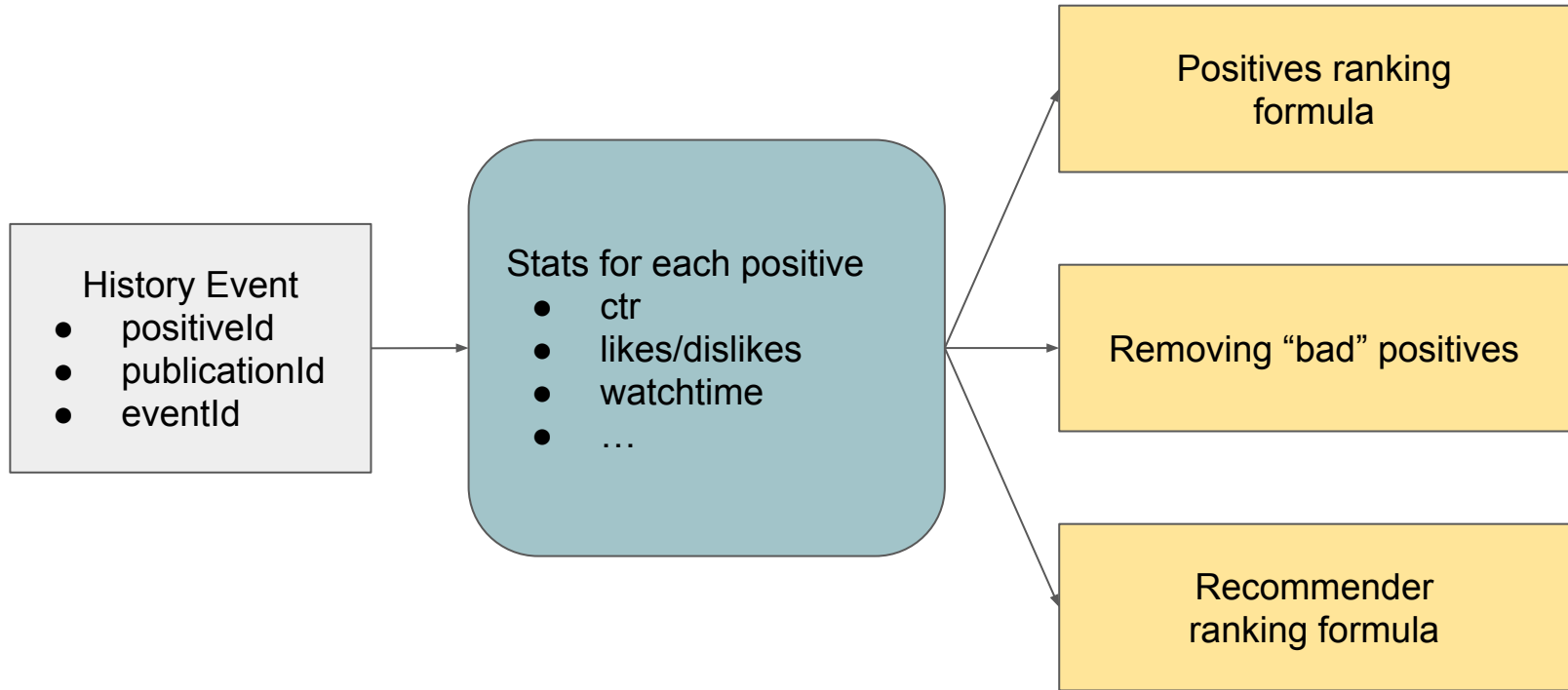
- prefer high weight points
- keep diversity in set



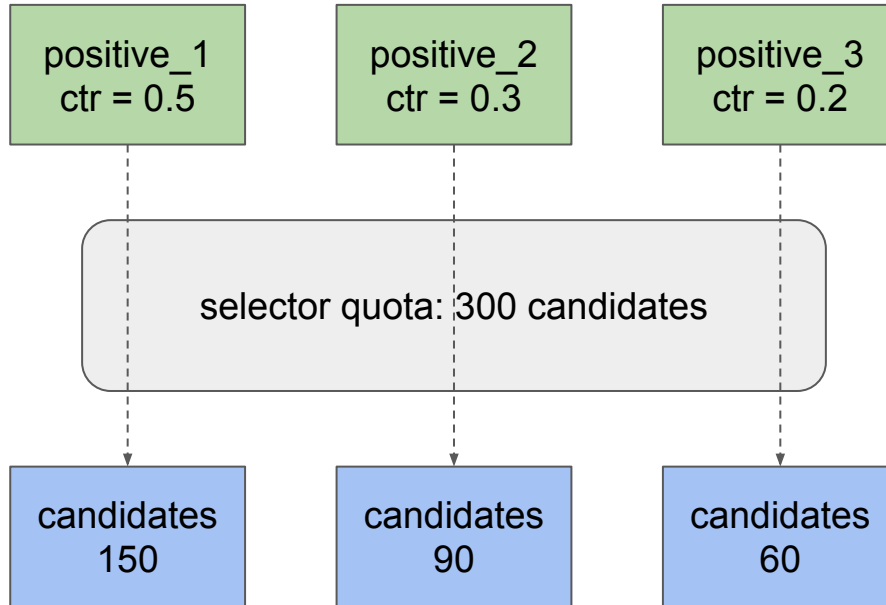
# Positives selection / Tree



# Positives selection / Formula



# Candidate selection / Bandits



# Conclusion

# Summary

1. Transparent recommendations pipeline
2. User has recommendation explanation
3. Feed relevance increase
4. Developers can trace, debug and improve system



Thank you!