Neural Networks for Recommendation Systems

Advanced ML with TensorFlow on GCP

End-to-End Lab on Structured Data ML

Production ML Systems

Image Classification Models

Sequence Models

Recommendation Systems

Learn how to...

Combine content-based, knowledge-based, and collaborative filtering recommendation systems

Use neural networks to make hybrid recommendation systems

Real-world recommendation systems are a hybrid of three broad theoretical approaches

Recommendation Systems

Content-Based

Recommend items based on content features.

- No need for data about other users
- Can recommend niche items

- Need domain knowledge
- Only safe recommendations

Collaborative Filtering

Based on user behavior only. Recommend items based on users with similar patterns.

- No domain knowledge
- Serendipity
- Great starting point

- Cold start, fresh items/ users
- Sparsity
- No context features

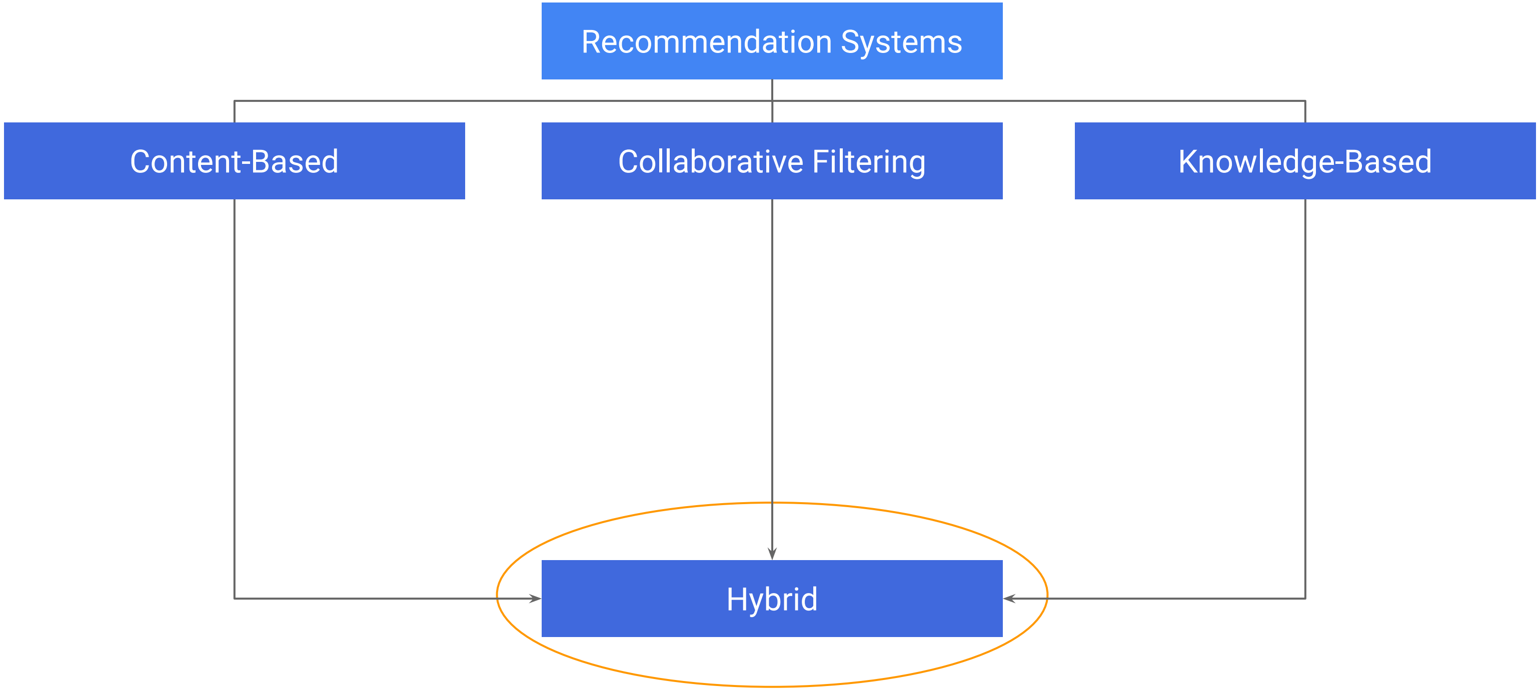
Knowledge-Based

Ask users for preferences.

- No interaction data needed
- Usually high-fidelity data from user self-reporting

- Need user data
- Need to be careful with privacy concerns

Real-world recommendation systems are a hybrid of three broad theoretical approaches



Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1 User ratings of item between 1 to 5 stars.

2 User reviews about experience with item.

3 User-answered questions about item.

4 Number of times user added item to cart.

A. Content-based

B. Collaborative filtering

C. Knowledge-based

D. A & B

E. A & C

F. B & C

G. A, B, & C

Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1. User ratings of item between 1 to 5 stars.
2. User reviews about experience with item.
3. User-answered questions about item.
4. Number of times user added item to cart.

A. Content-based

B. Collaborative filtering

C. Knowledge-based

D. A & B

E. A & C

F. B & C

G. A, B, & C

Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1. User ratings of item between 1 to 5 stars.
2. User reviews about experience with item.
3. User-answered questions about item.
4. Number of times user added item to cart.

A. Content-based
B. Collaborative filtering
C. Knowledge-based
D. A & B
E. A & C
F. B & C
G. A, B, & C

Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1. User ratings of item between 1 to 5 stars.
2. User reviews about experience with item.
3. User-answered questions about item.
4. Number of times user added item to cart.

A. Content-based
B. Collaborative filtering
C. Knowledge-based
D. A & B
E. A & C
F. B & C
G. A, B, & C

Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1. User ratings of item between 1 to 5 stars.
2. User reviews about experience with item.
3. User-answered questions about item.
4. Number of times user added item to cart.

A. Content-based

B. Collaborative filtering

C. Knowledge-based

D. A & B

E. A & C

F. B & C

G. A, B, & C

Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1. User ratings of item between 1 to 5 stars.
2. User reviews about experience with item.
3. User-answered questions about item.
4. Number of times user added item to cart.

A. Content-based

B. Collaborative filtering

C. Knowledge-based

D. A & B

E. A & C

F. B & C

G. A, B, & C

Quiz: Popular recommendation systems

If we have ONLY the following data to recommend items for users to buy, what type of recommendation system should we use?

1. User ratings of item between 1 to 5 stars.
2. User reviews about experience with item.
3. User-answered questions about item.
4. Number of times user added item to cart.

A. Content-based

B. Collaborative filtering

C. Knowledge-based

D. A & B

E. A & C

F. B & C

G. A, B, & C

Lab

Design a hybrid recommendation system

In this lab, you will learn how to design a hybrid recommendation system to recommend movies to users.

Lab Steps

1. Think of datasets we can use for each type:
   Content-based
   Collaborative Filtering
   Knowledge-based

2. Structured and unstructured

3. Explicit and implicit feedback

Content-based recommendation models

1 Structured

- Genres
- Themes
- Actors/directors involved
- Professional ratings

2 Unstructured

- Movie summary text
- Stills from movie
- Movie trailer
- Professional reviews

Collaborative filtering

1 Structured

User ratings

User views

User wishlist/cart history

User purchase/return history

2 Unstructured

User reviews

User-answered questions

User-submitted photos

User-submitted videos

Knowledge-based

1 Structured

Demographic information

Location/country/language

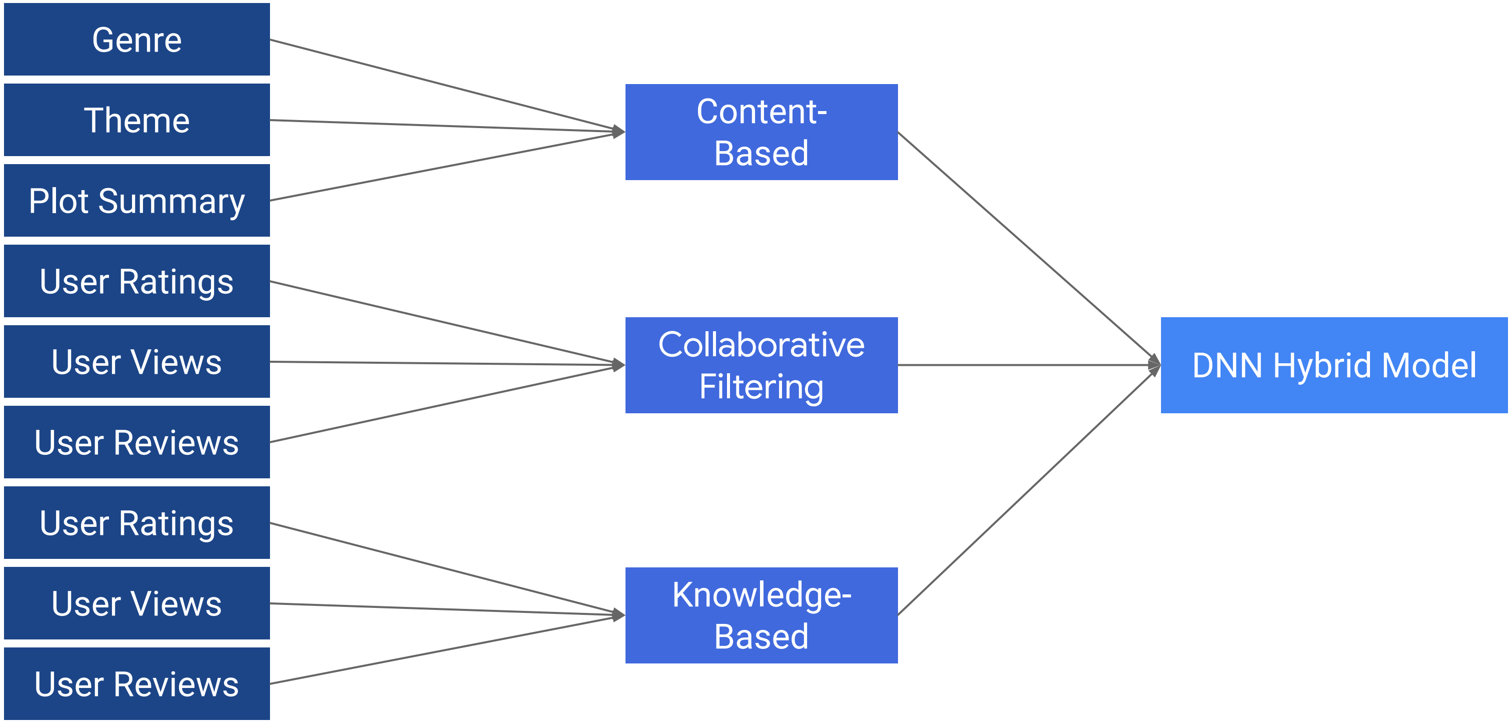
Genre preferences

Global filters

2 Unstructured

User “about me” snippets

Hybrid model



Deep learning for product recommendations

1. No new concepts (it's just a structured data model).
2. There's a lot of data to bring together.
3. Need multiple ML models (an ML pipeline).

Quiz: Hybrid recommendation systems

What is important to have when making a hybrid recommendation system?

A. Data collection with recommendation in mind
B. Many different datasets
C. More than one recommendation model type
D. An ML model pipeline
E. All of the above

Quiz: Hybrid recommendation systems

What is important to have when making a hybrid recommendation system?

A. Data collection with recommendation in mind
B. Many different datasets
C. More than one recommendation model type
D. An ML model pipeline
E. All of the above

Lab

Neural network hybrid recommendation system with Google Analytics data

In this lab, you will learn how to complete TODOs in hybrid_recommendations.ipynb.

Lab Steps

1. Create input layer for feature columns.

2. Create neural network layers.

3. Create output layer with our labels.

Adding context

1. An item is not just an item.
2. A user is not just a user.
3. The context it is experienced in changes perception.
4. This affects sentiment.

Context components

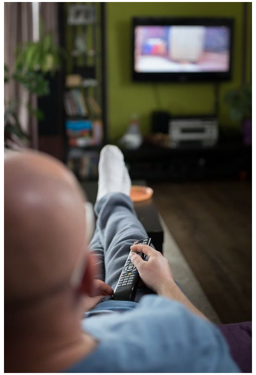
Mood at the time?

Who else experiencing item with?

Where experiencing item?

When experiencing item?

Special occasion?



Context components

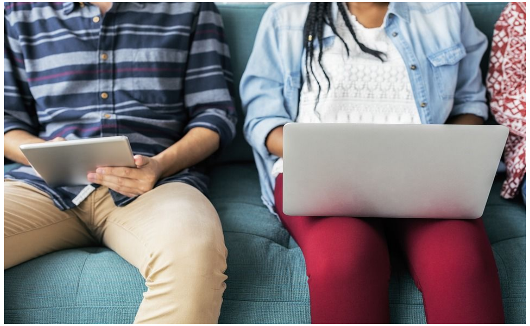
Mood at the time?

Who else experiencing item with?

Where experiencing item?

When experiencing item?

Special occasion?



Context components

Mood at the time?

Who else experiencing item with?

Where experiencing item?

When experiencing item?

Special occasion?

Context components

Mood at the time?

Who else experiencing item with?

Where experiencing item?

When experiencing item?

Special occasion?

Context components

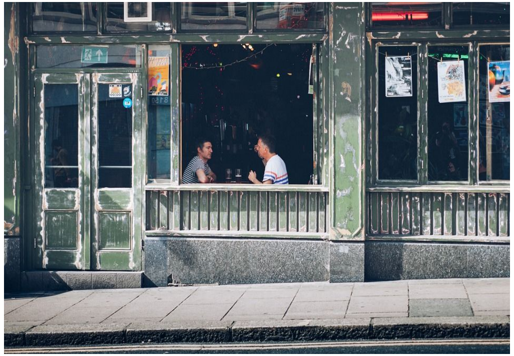
Mood at the time?

Who else experiencing item with?

Where experiencing item?

When experiencing item?

Special occasion?



Context-aware recommendation systems (CARS)

Traditional CF RS:

Users x Items → Ratings

Contextual CF RS:

Users x Items x Contexts → Ratings

User-item-context example data

User	Item	Who	Where	When	Rating
U1	M1	Kids	Home	Weekend	5
U1	M2	Family	Theater	Weekend	4
U1	M3	Partner	Event	Weekday	5
U2	M1	Friends	Home	Weekend	3
U2	M2	Family	Home	Weekday	4
U3	M2	Kids	Theater	Weekday	2
U3	M3	Partner	Home	Weekend	1
U2	M3	Partner	Home	Weekday	$\boxed{?}$

User-item-context example data

User	Item	Who	Where	When	Rating
U1	M1	Kids	Home	Weekend	5
U1	M2	Family	Theater	Weekend	4
U1	M3	Partner	Event	Weekday	5
U2	M1	Friends	Home	Weekend	3
U2	M2	Family	Home	Weekday	4
U3	M2	Kids	Theater	Weekday	2
U3	M3	Partner	Home	Weekend	1
U2	M3	Partner	Home	Weekday	?

User-item-context example data

User	Item	Who	Where	When	Rating
U1	M1	Kids	Home	Weekend	5
U1	M2	Family	Theater	Weekend	4
U1	M3	Partner	Event	Weekday	5
U2	M1	Friends	Home	Weekend	3
U2	M2	Family	Home	Weekday	4
U3	M2	Kids	Theater	Weekday	2
U3	M3	Partner	Home	Weekend	1
U2	M3	Partner	Home	Weekday	?

Quiz: The effect of context on user-item interactions

We learned that context can be important when thinking about user-item interactions. Which is NOT an example of context for a user-item interaction?

A. Watching a movie at the theater
B. Watching a movie with family
C. Relaxing on the weekend with a movie
D. The length of a movie
E. Your mood while watching a movie
F. Watching a movie late at night
G. Watching a movie at a big watch party

Quiz: The effect of context on user-item interactions

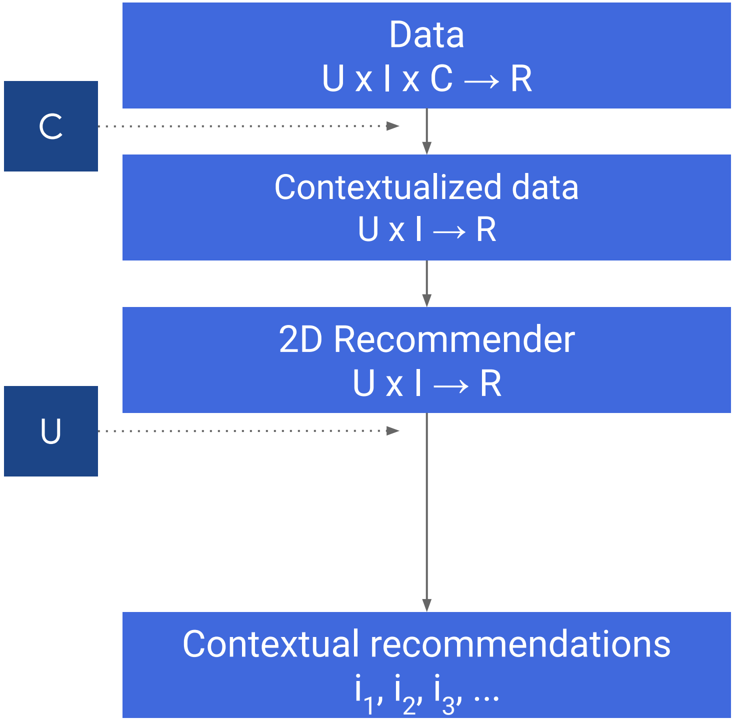
We learned that context can be important when thinking about user-item interactions. Which is NOT an example of context for a user-item interaction?

A. Watching a movie at the theater
B. Watching a movie with family
C. Relaxing on the weekend with a movie
D. The length of a movie
E. Your mood while watching a movie
F. Watching a movie late at night
G. Watching a movie at a big watch party

There are three main types of context-aware recommendation systems, or CARS, algorithms

1. Contextual prefiltering
2. Contextual postfiltering
3. Contextual modeling

Contextual prefiltering



Contextual prefiltering

- Reduction-Based Approach, 2005
- Exact and Generalized Prefiltering, 2009
- Item Splitting, 2009
- User Splitting, 2011
- Dimension as Virtual Items, 2011
- User-Item Splitting, 2014

Item splitting

User	Item	Time	Rating
U1	M1	Weekend	5
U2	M1	Weekend	5
U3	M1	Weekend	4
U4	M1	Weekend	5
U1	M1	Weekday	2
U2	M1	Weekday	3
U3	M1	Weekday	2
U4	M1	Weekday	2

Item splitting

User	Item	Time	Rating
U1	M1	Weekend	5
U2	M1	Weekend	5
U3	M1	Weekend	4
U4	M1	Weekend	5
U1	M1	Weekday	2
U2	M1	Weekday	3
U3	M1	Weekday	2
U4	M1	Weekday	2

Item splitting

User	Item	Rating
U1	M1,1	5
U2	M1,1	5
U3	M1,1	4
U4	M1,1	5
U1	M1,2	2
U2	M1,2	3
U3	M1,2	2
U4	M1,2	2

$$t_{mean} = \left| \frac{\mu_{i_c} - \mu_{i_{\bar{c}}}}{\sqrt{s_{i_c}/n_{i_c} + s_{i_{\bar{c}}}/n_{i_{\bar{c}}}}} \right|$$

User splitting

User	Item	Time	Rating
U1	M1	Weekend	5
U1	M1	Weekday	2
U2	M1	Weekend	5
U2	M1	Weekday	3
U3	M1	Weekend	4
U3	M1	Weekday	2
U4	M1	Weekend	5
U4	M1	Weekday	2

User splitting

User	Item	Rating
U1,1	M1	5
U1,2	M1	2
U2,1	M1	5
U2,2	M1	3
U3,1	M1	4
U3,2	M1	2
U4,1	M1	5
U4,2	M1	2

User-item splitting

User	Item	Time	Location	Rating
U1	M1	Weekend	Home	5
U1	M1	Weekday	Theater	2
U2	M1	Weekend	Theater	5
U2	M1	Weekday	Home	3
U3	M1	Weekend	Home	4
U3	M1	Weekday	Theater	2
U4	M1	Weekend	Theater	5
U4	M1	Weekday	Home	2

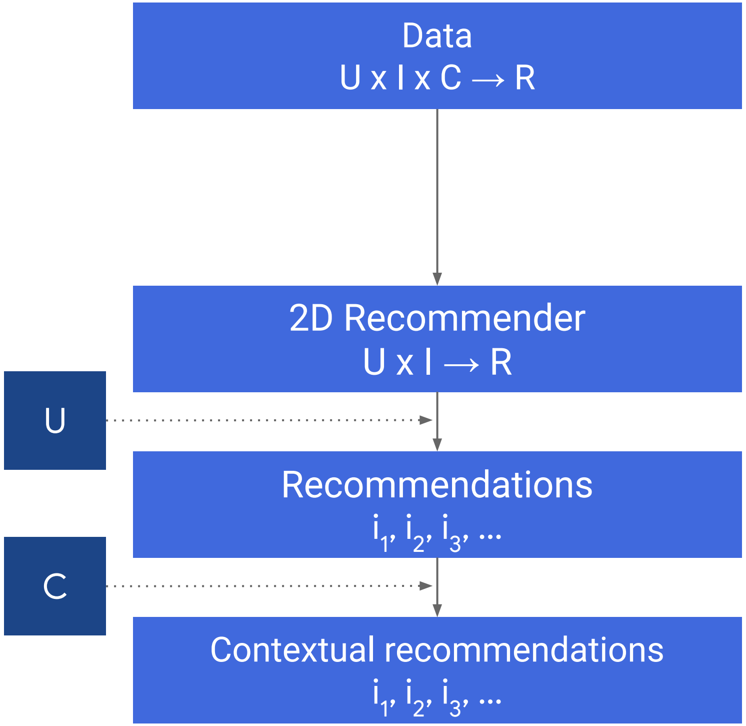
User-item splitting

User	Item	Time	Location	Rating
U1,1	M1,1	Weekend	Home	5
U1,2	M1,2	Weekday	Theater	2
U2,1	M1,2	Weekend	Theater	5
U2,2	M1,1	Weekday	Home	3
U3,1	M1,1	Weekend	Home	4
U3,2	M1,2	Weekday	Theater	2
U4,1	M1,2	Weekend	Theater	5
U4,2	M1,1	Weekday	Home	2

User-item splitting

User	Item	Rating
U1,1	M1,1	5
U1,2	M1,2	2
U2,1	M1,2	5
U2,2	M1,1	3
U3,1	M1,1	4
U3,2	M1,2	2
U4,1	M1,2	5
U4,2	M1,1	2

Contextual postfiltering



Contextual postfiltering

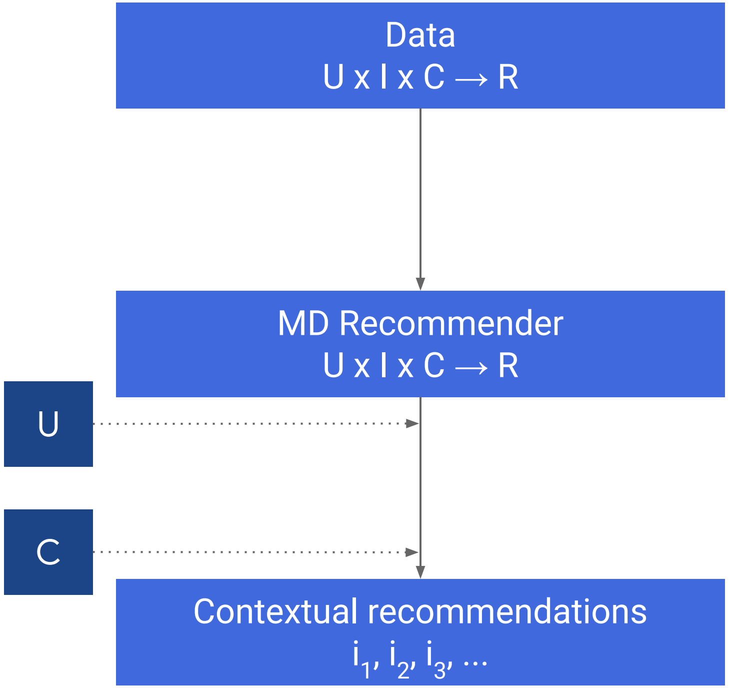
Weight, 2009

$$r'_{ij} = r_{ij} * P$$

Filter, 2009

$$P < P_*$$

Contextual modeling



Contextual modeling

- Tensor Factorization, 2010
- Factorization Machines, 2011
- Deviation-Based Context-Aware Matrix Factorization, 2011
- Deviation-Based Sparse Linear Method, 2014
- Similarity-Based Context-Aware Matrix Factorization, 2015
- Similarity-Based Sparse Linear Method, 2015

Deviation-based context-aware matrix factorization

- How is user's rating deviated?
- Contextual rating deviation (CRD)
- Looks at the deviations of users across context dimensions

Deviation-based context-aware matrix factorization

Context	Location	Time	Who
C1	Home	Weekend	Family
C2	Home	Weekend	Friend
C3	Home	Weekday	Family
C4	Home	Weekday	Friend
C5	Theater	Weekend	Family
C6	Theater	Weekend	Friend
C7	Theater	Weekday	Family
C8	Theater	Weekday	Friend

Deviation-based context-aware matrix factorization

Context	Location	Time	Who
C1	Home	Weekend	Family
C8	Theater	Weekday	Friend
CRD(Dim)	0.8	-0.2	0.1

Deviation-based context-aware matrix factorization

Biased matrix factorization in traditional RS

Global average rating

User-item interaction

$$\hat{r}_{ui} = \mu + b_u + b_i + p_u^T q_i$$

User bias

Item bias

Deviation-based context-aware matrix factorization

CAMF_C approach

Global average rating

User-item interaction

$$\hat{r}_{uic_1c_2 \ldots c_N} = \mu + b_u + b_i + p_u^T q_i + \sum_{j=1}^{N} CRD(c_j)$$

User bias

Item bias

Contextual Rating

Contextual Rating Deviation

Deviation-based context-aware matrix factorization

CAMF_CU approach

$$\widehat{r}_{uic_1c_2...c_N} = \mu + \sum_{j=1}^{N} CRD(c_j, u) + b_i + p_u^T q_i$$

CAMF_CI approach

$$\widehat{r}_{uic_1c_2...c_N} = \mu + b_u + \sum_{j=1}^{N} CRD(c_j, i) + p_u^T q_i$$

Quiz: Context-aware recommendation system algorithms

Which context-aware recommendation system type produces non-contextual recommendations that it later adjusts via context into contextual recommendations?

A. Contextual modeling
B. Contextual postprocessing
C. Contextual prefiltering
D. Contextual adjustment
E. Contextual postfiltering
F. Contextual aggregation
G. None of the above

Quiz: Context-aware recommendation system algorithms

Which context-aware recommendation system type produces non-contextual recommendations that it later adjusts via context into contextual recommendations?

A. Contextual modeling

B. Contextual postprocessing

C. Contextual prefiltering

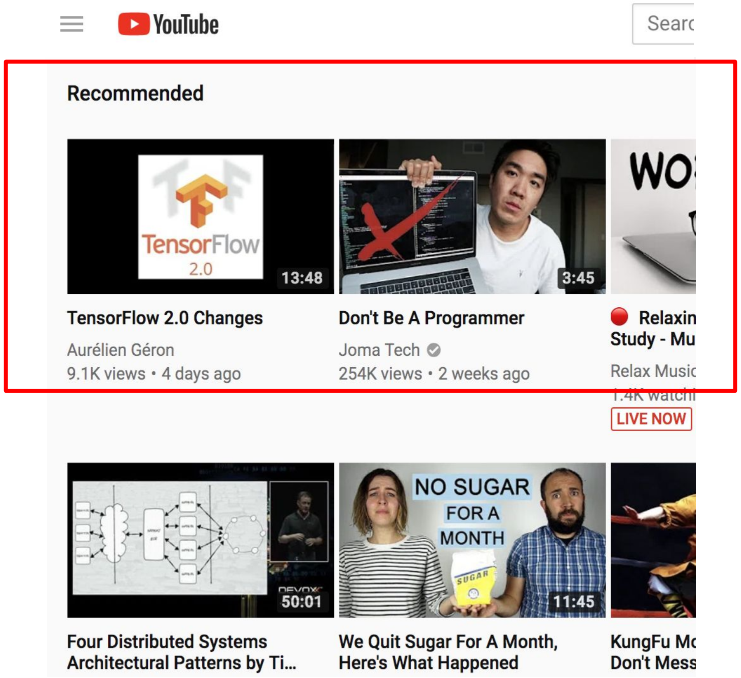
D. Contextual adjustment

E. Contextual postfiltering

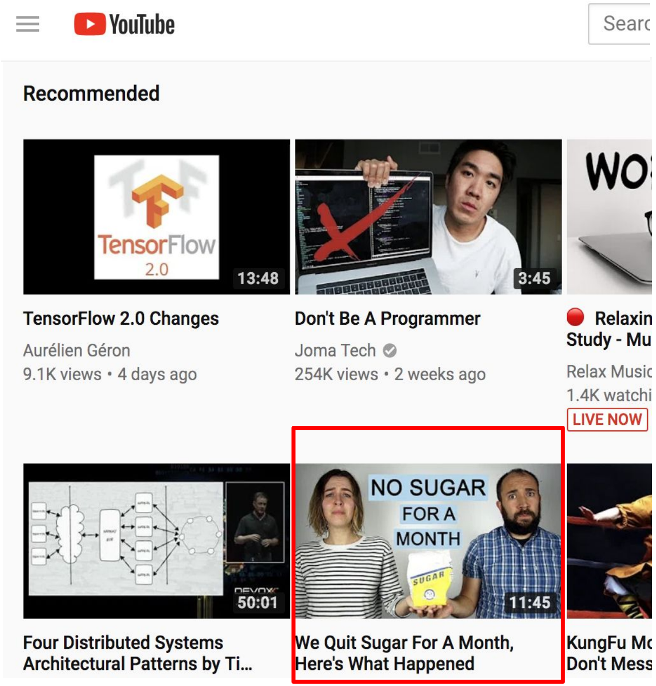
F. Contextual aggregation

G. None of the above

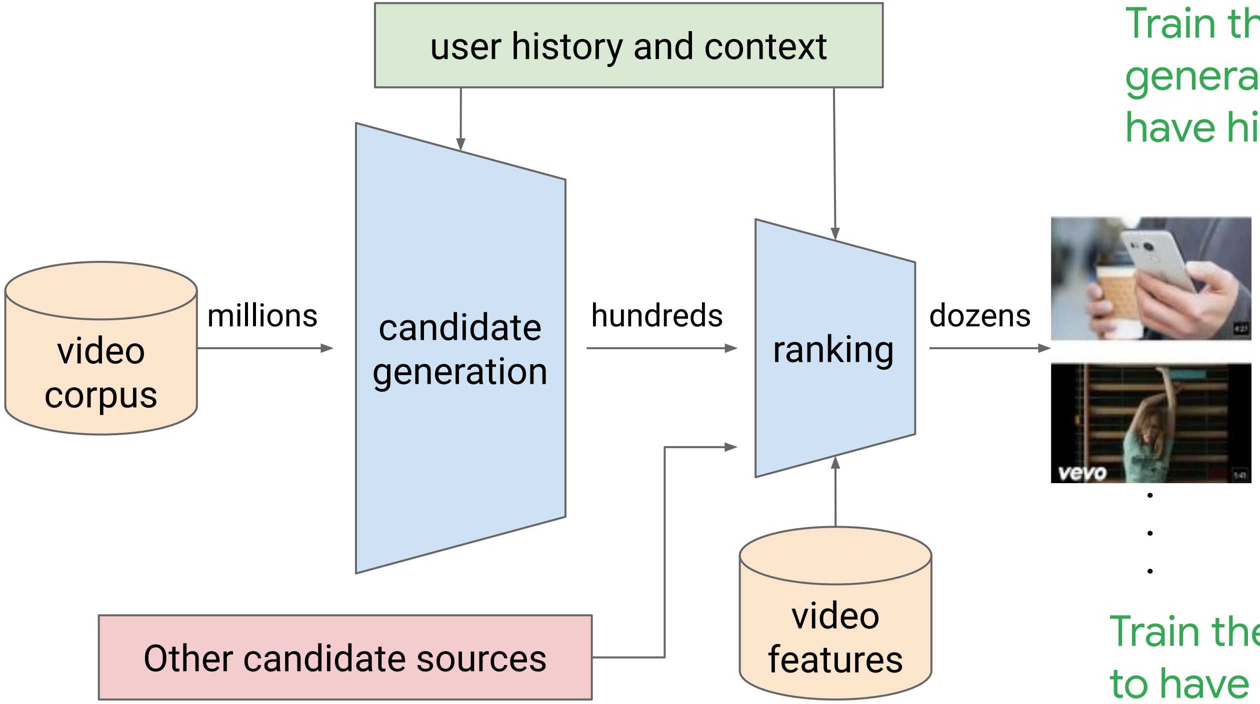
Recommendation engines identify things that a user may like based on what they've watched in the past



Recommendation engines suggest new items that a user might not have thought to search for



YouTube uses two neural networks to recommend videos



Train the candidate generation NN to have high precision.

Train the ranking NN to have high recall.

Quiz: Hybrid recommendation systems

YouTube uses two neural networks connected in an ML pipeline. What metric is the candidate generation network trained to maximize? What metric is the ranking network trained to maximize?

A. Precision, Similarity

B. Similarity, Recall

C. Recall, Precision

D. F1 Score, Precision

E. AUC, Recall

F. Precision, Recall

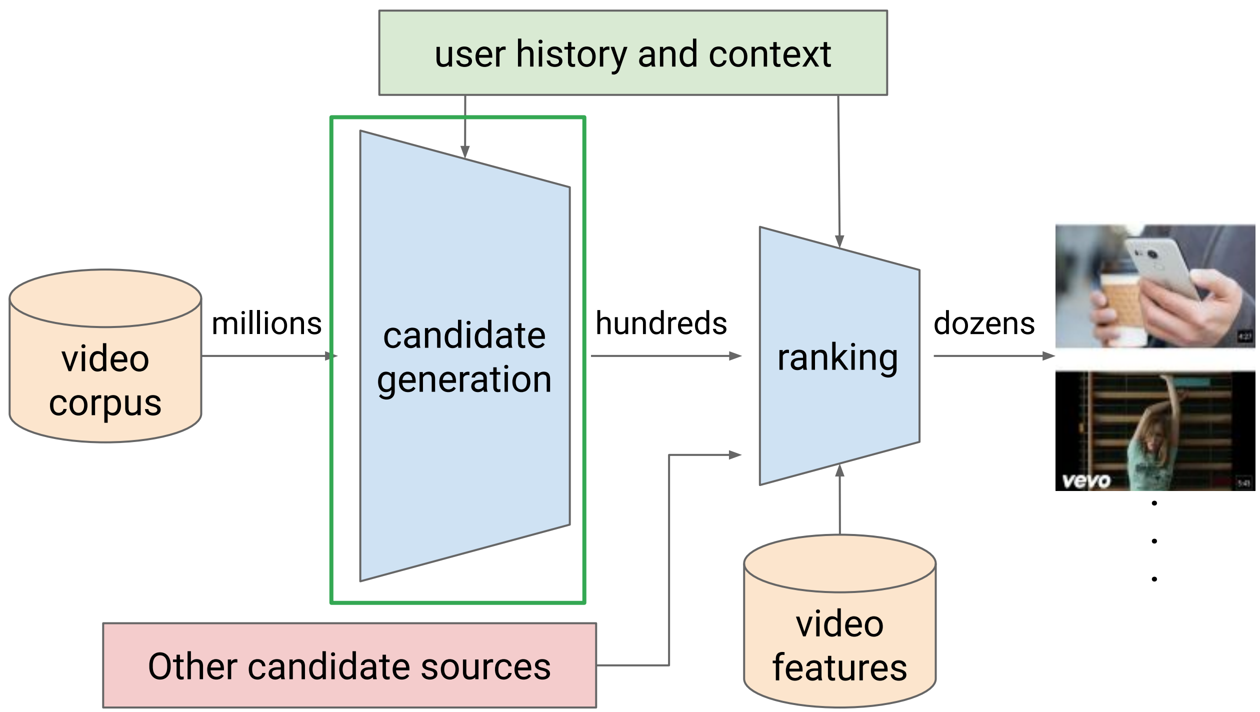
G. None of the above

Quiz: Hybrid recommendation systems

YouTube uses two neural networks connected in an ML pipeline. What metric is the candidate generation network trained to maximize? What metric is the ranking network trained to maximize?

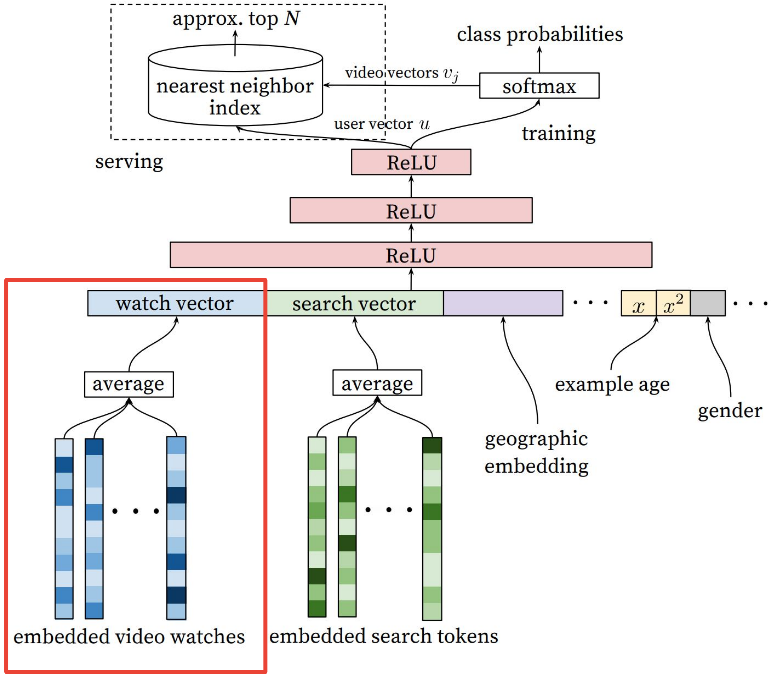
A. Precision, Similarity
B. Similarity, Recall
C. Recall, Precision
D. F1 Score, Precision
E. AUC, Recall
F. Precision, Recall
G. None of the above

Candidate generation



Candidate generation consists of intelligently assembling many other ML models (1/3)

1. Get item embedding from, e.g., WALS.
2. Find last 10 videos watched by user.
3. Average the embeddings of those 10 videos.
4. This is the watch vector.

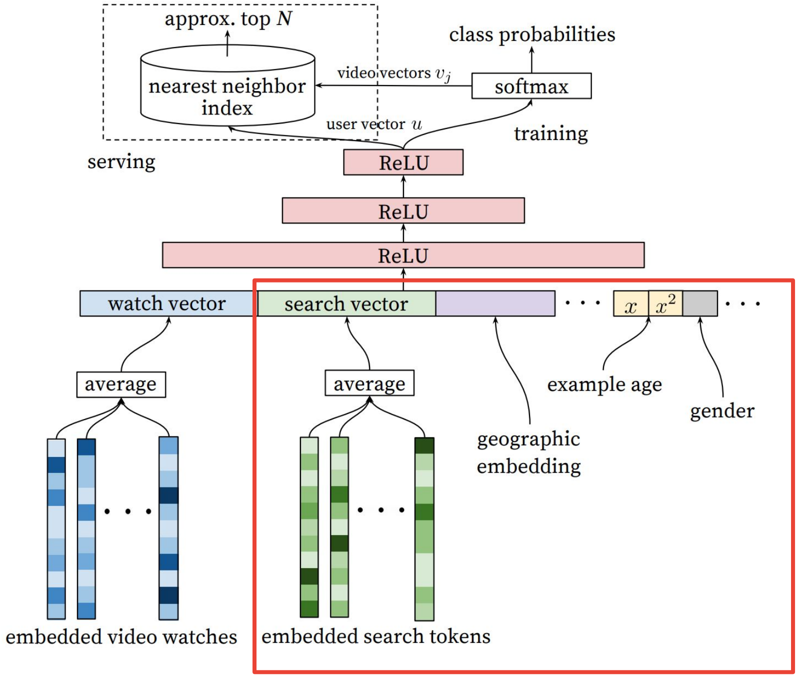


Candidate generation consists of intelligently assembling many other ML models (2/3)

5. Do the same thing with past search queries (collaborative filtering for next search term).

6. Add knowledge about user (e.g., location, gender).

7. Add example age to avoid overemphasizing older videos.

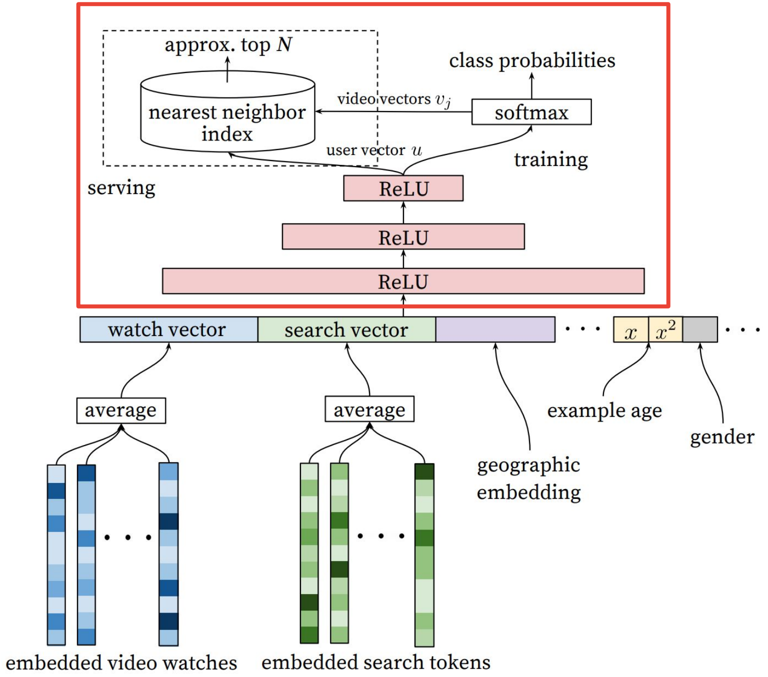


Candidate generation consists of intelligently assembling many other ML models (3/3)

8. Train a DNN Classifier.

9. Treat the last-but-one layer as a user embedding.

10. Use output of DNN Classifier and user embedding to generate candidates.



Quiz: Candidate generation network

During training of the candidate generation network, what output layer should we be using, and what should we be predicting with it?

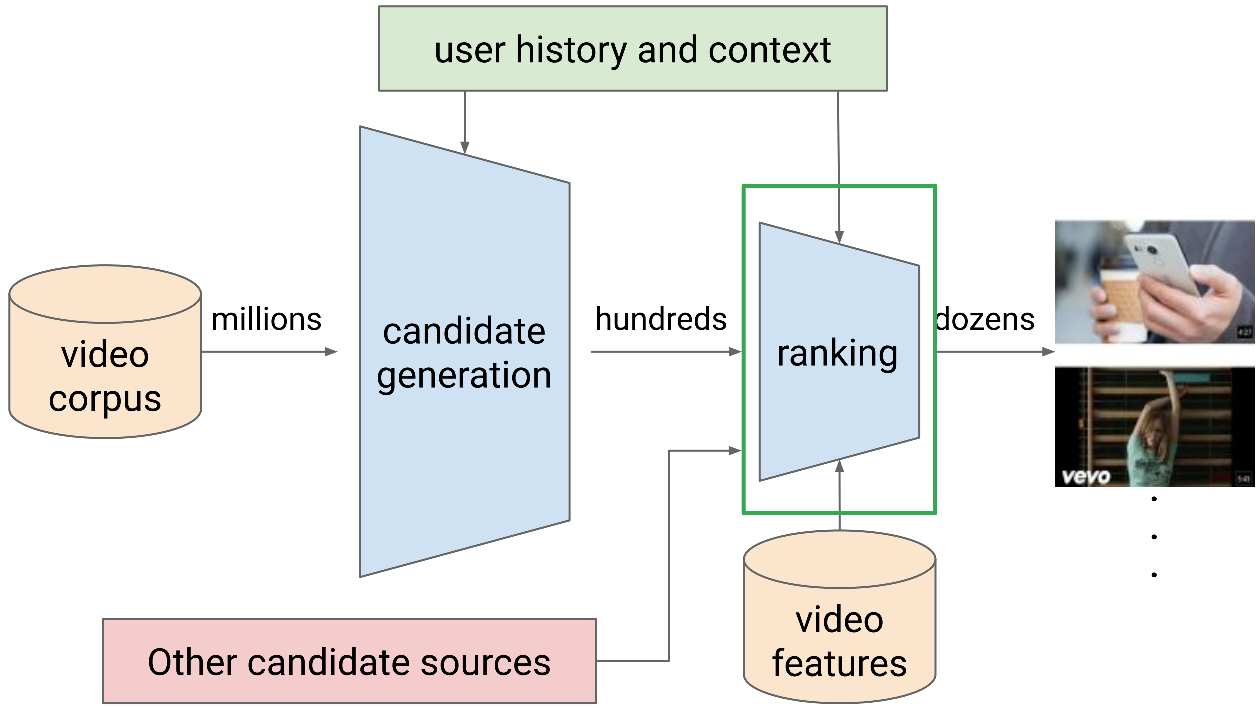
A. Linear, Embeddings
B. Linear, Probabilities
C. Linear, Unbounded real numbers
D. Softmax, Embeddings
E. Softmax, Probabilities
F. Softmax, Unbounded real numbers
G. None of the above

Quiz: Candidate generation network

During training of the candidate generation network, what output layer should we be using, and what should we be predicting with it?

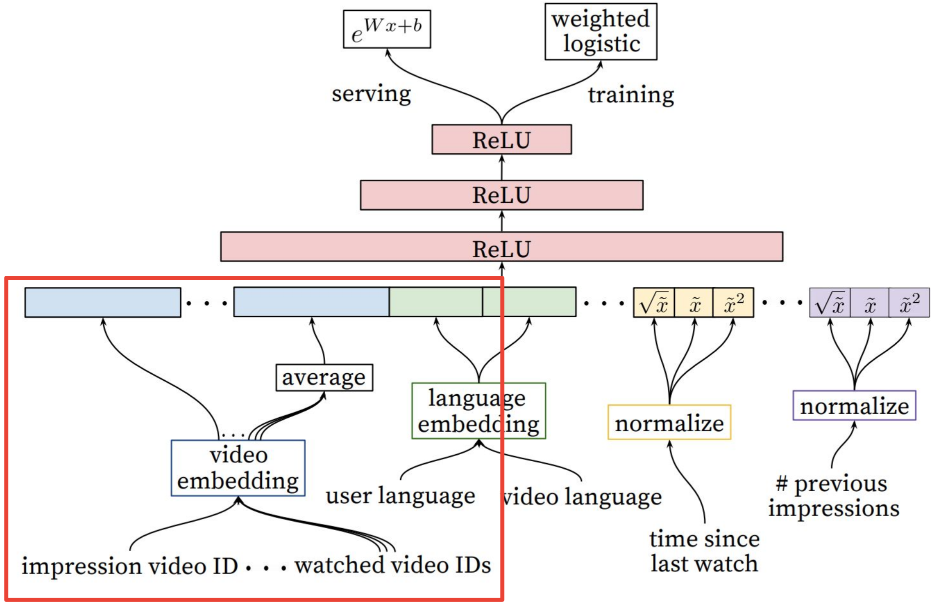
A. Linear, Embeddings
B. Linear, Probabilities
C. Linear, Unbounded real numbers
D. Softmax, Embeddings
E. Softmax, Probabilities
F. Softmax, Unbounded real numbers
G. None of the above

Ranking


user history and context
video
corpus
millions
candidate
generation
hundreds
ranking
dozens
vevo
Other candidate sources
video
features

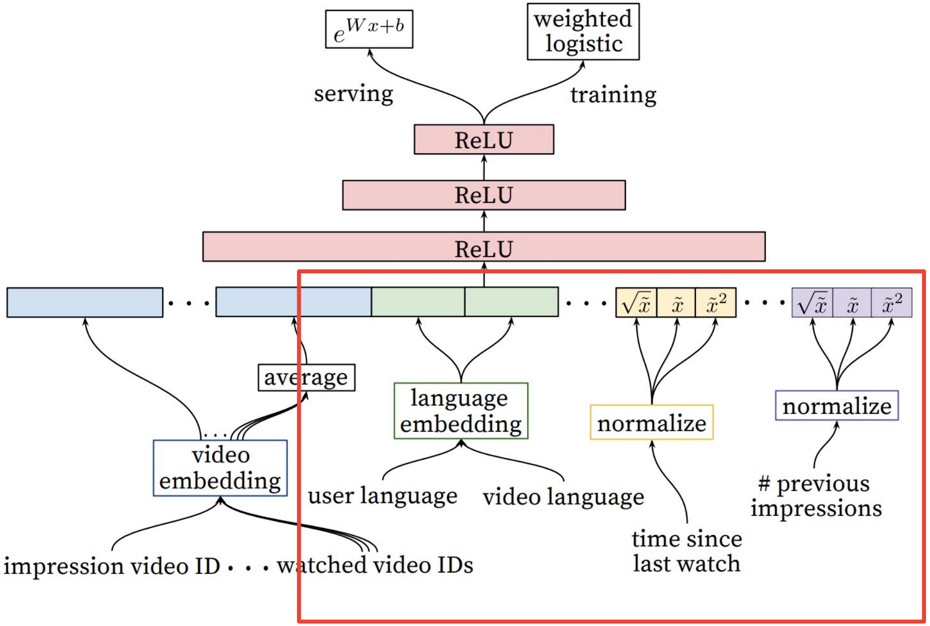
The ranking network uses more tailored features (1/3)

1. Videos suggested to user.
2. Videos watched by user.
3. Both individual and average embedding.



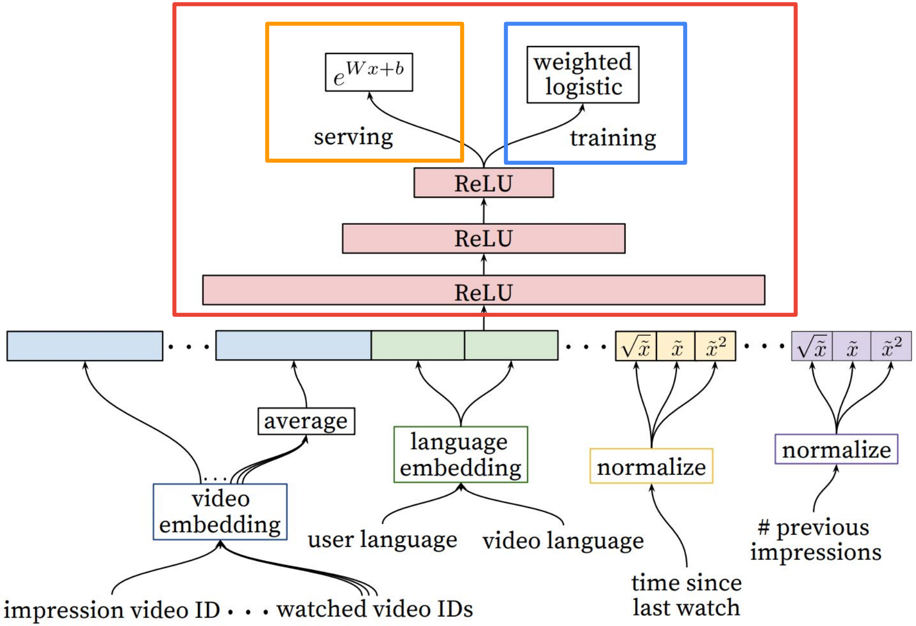
The ranking network uses more tailored features (2/3)

4. Hundreds of features including language embeddings and user behavior.



The ranking network uses more tailored features (3/3)

5. DNN Classifier whose output is used for ranking.



Quiz: Ranking neural network

For the ranking neural network, why do we use a weighted logistic for training?

A. The loss function is more numerically stable
B. Positive and negative impression weights differ
C. The probabilities need special formatting
D. Reduces user noise better

Quiz: Ranking neural network

For the ranking neural network, why do we use a weighted logistic for training?

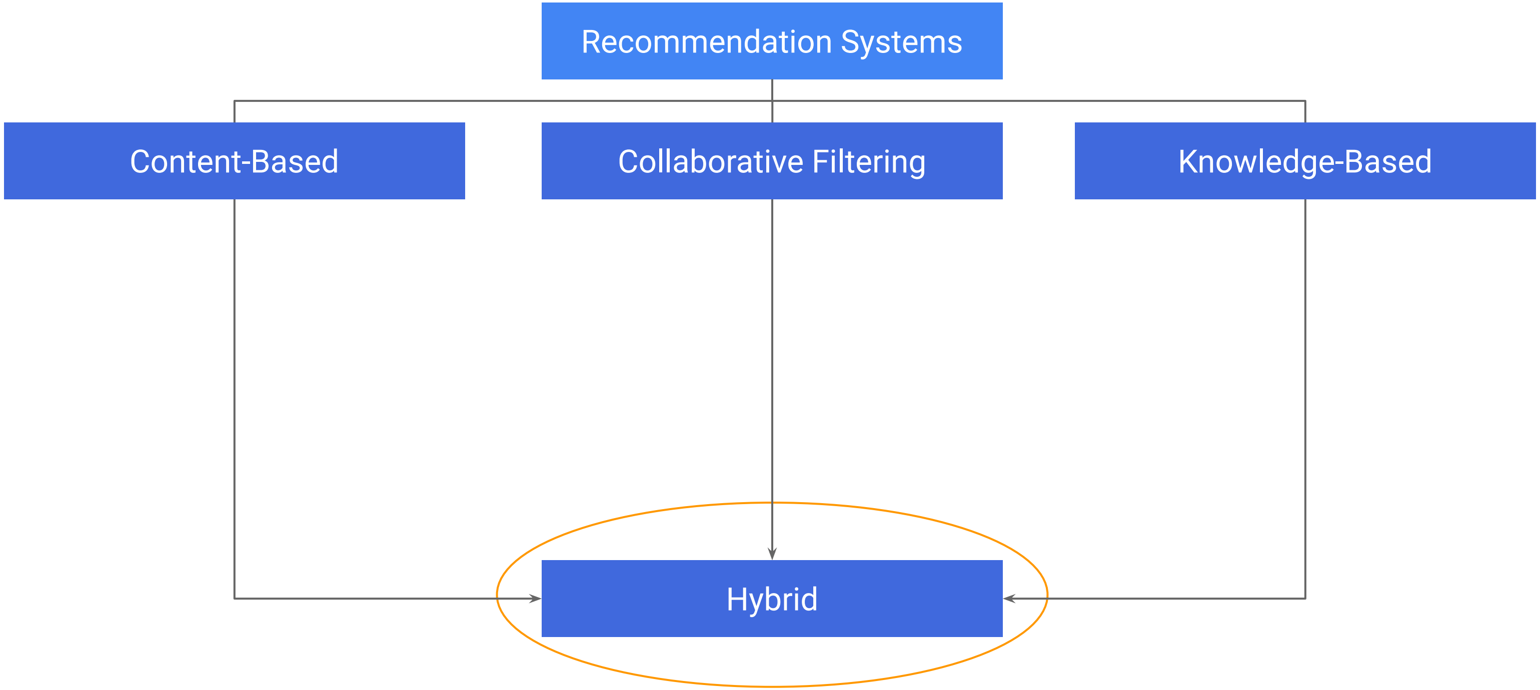
A. The loss function is more numerically stable

B. $\boxed{\text{Positive and negative impression weights differ}}$

C. The probabilities need special formatting

D. Reduces user noise better

Real-world recommendation systems are a hybrid of three broad theoretical approaches

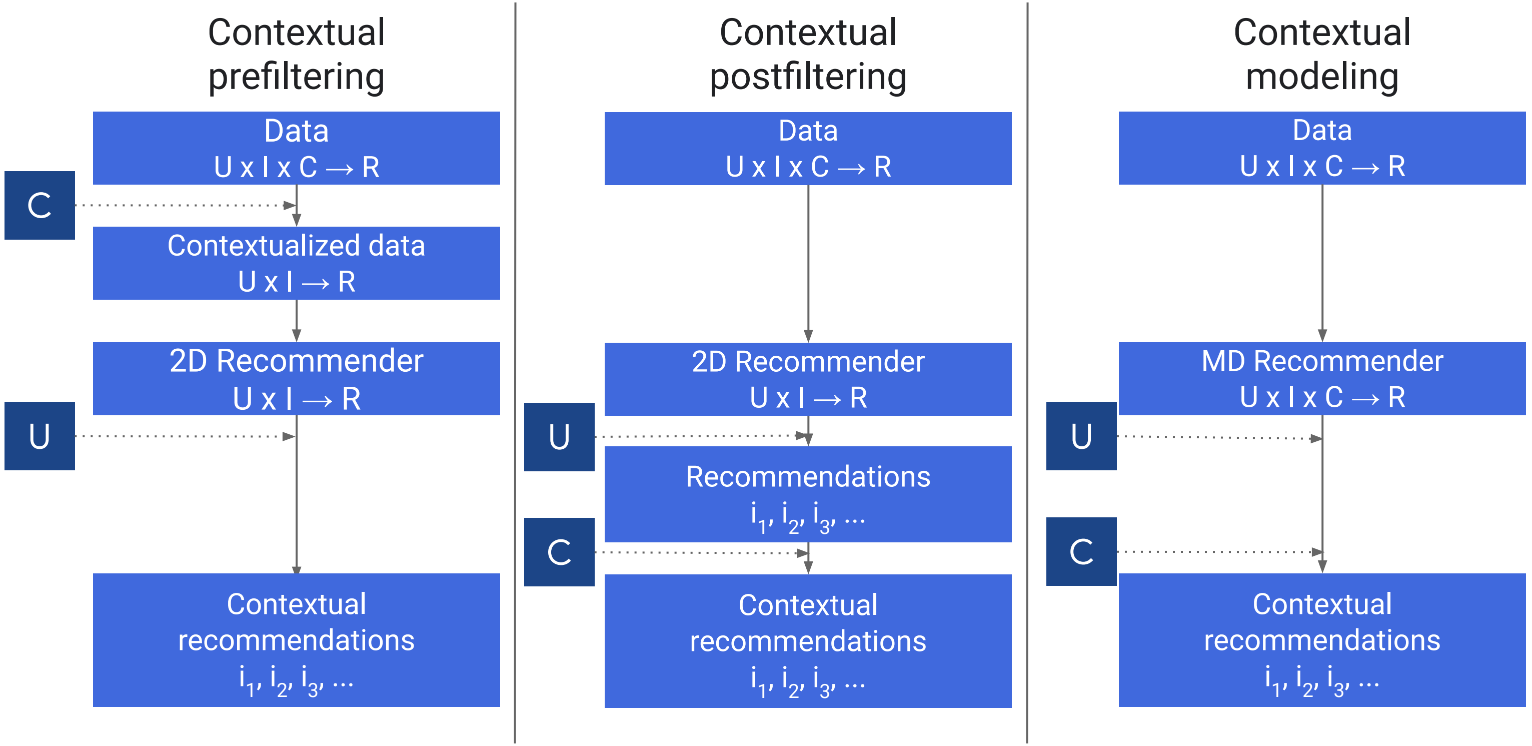


Context components

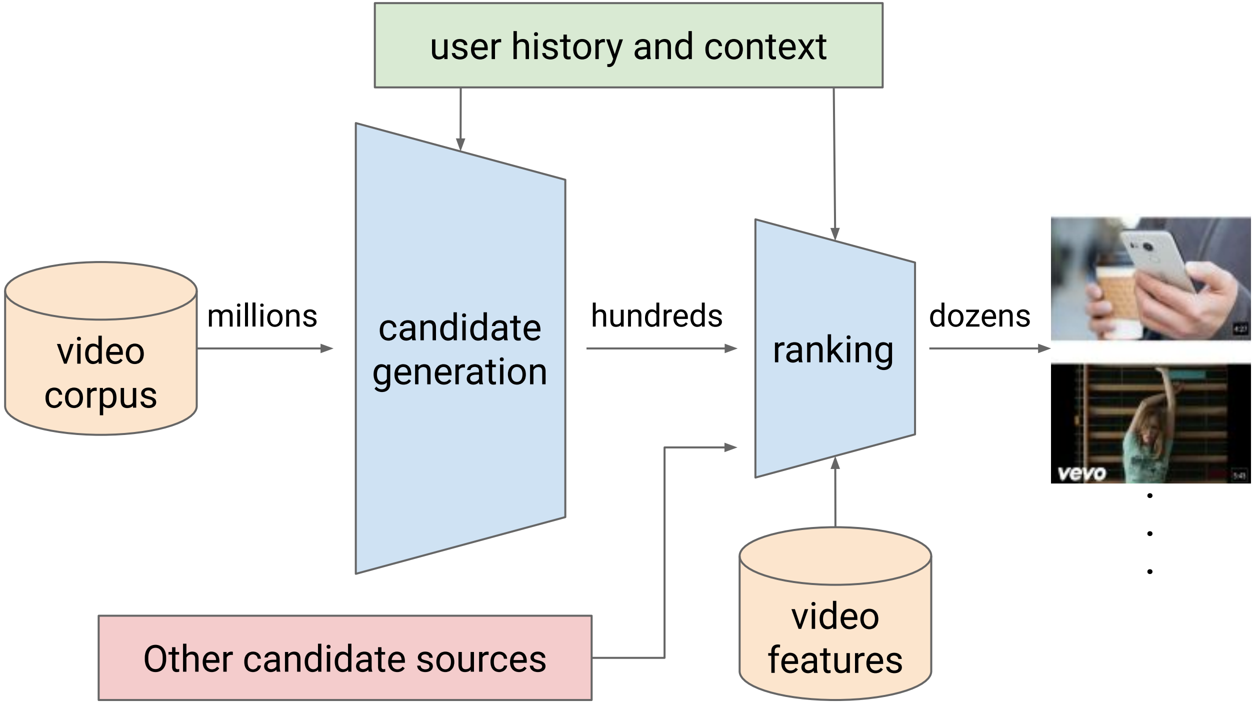




Context-aware recommendation systems



YouTube video recommendations



cloud.google.com