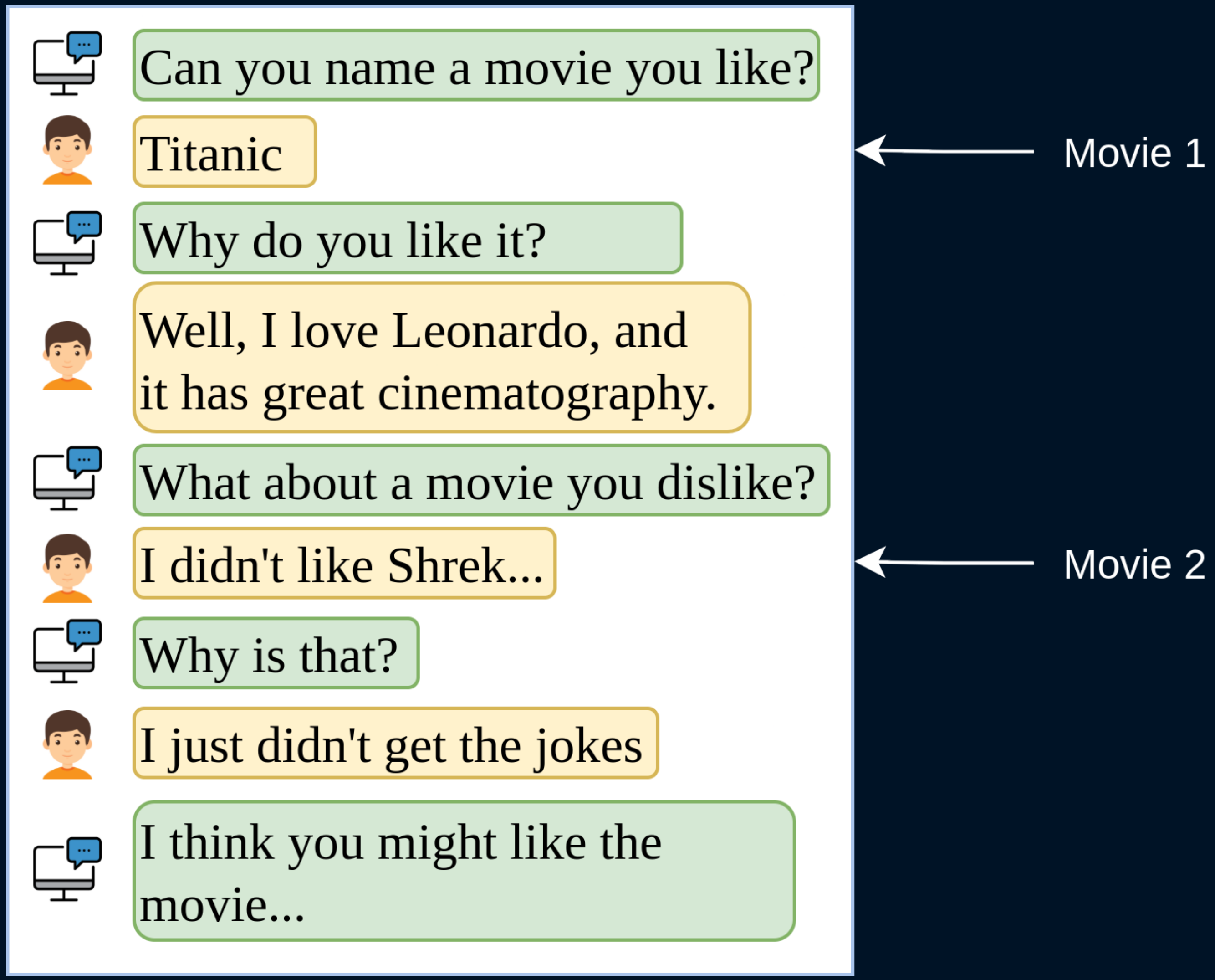
Towards Predicting Movie Preferences from Conversational Interactions

Sergey Volokhin, Joyce Ho, Oleg Rochlenko, Eugene Agichtein

Motivation

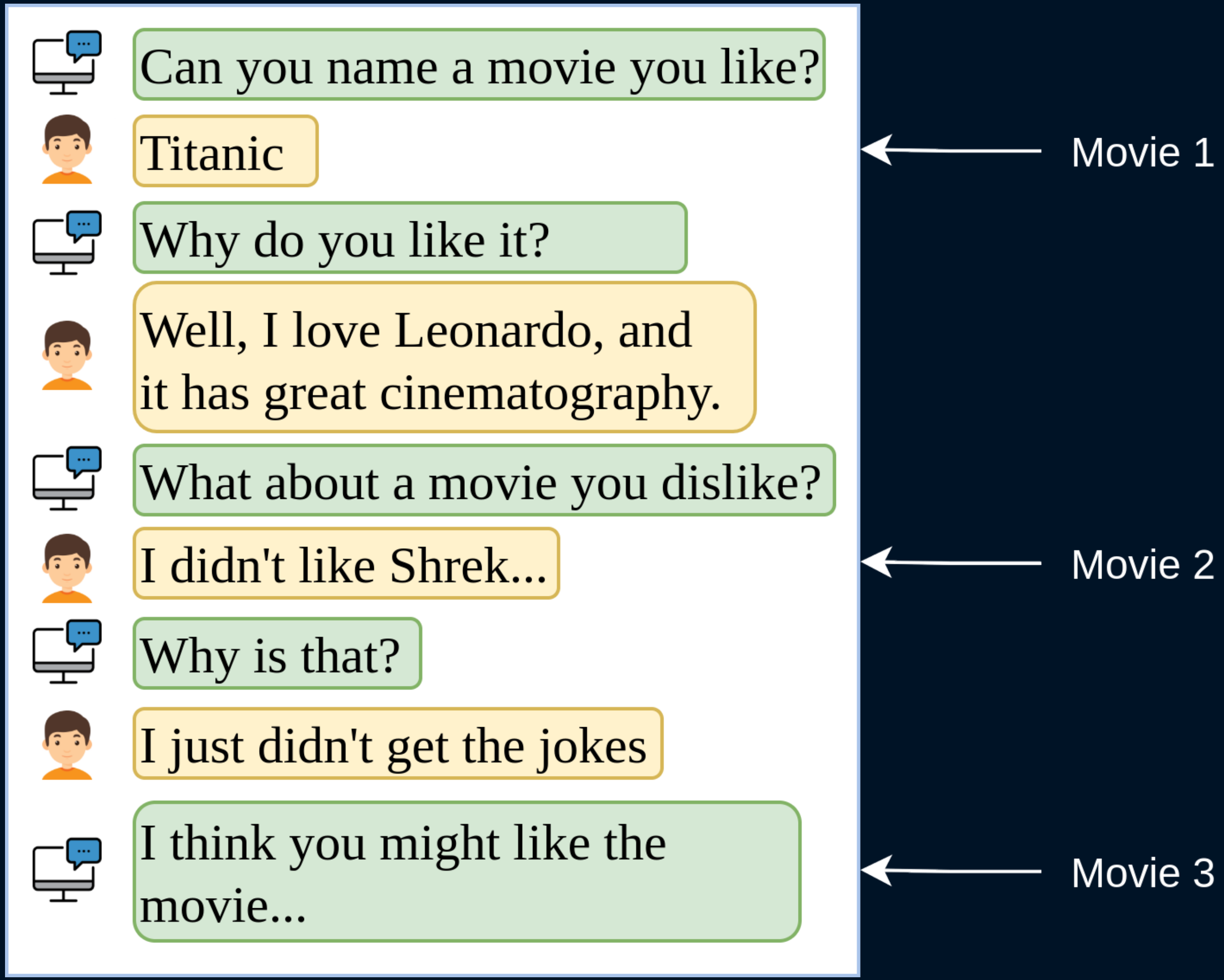
- Establishing user's preferences through a conversation for an effective recommendation remains an open question
- There exists little conversational data for such a task.

Problem



Can you name a movie you like?
Titanic
Why do you like it?
Well, I love Leonardo, and it has great cinematography.
What about a movie you dislike?
I didn't like Shrek...
Why is that?
I just didn't get the jokes
I think you might like the movie...
Movie 1
Movie 2

Problem Statement



Contributions

- Development of a public conversational dataset ***MovieSent***, annotated with:

Contributions

- Development of a public conversational dataset ***MovieSent***, annotated with:
  - Entities' IDs
  - Fine-grained user sentiment

Contributions

- Development of a public conversational dataset ***MovieSent***
- A new conversational recommendation method ***"Conversational Collaborative Filtering using External Data"***

Contributions

- Development of a public conversational dataset ***MovieSent***
- A new conversational recommendation method ***ConvExtr***, which:

Contributions

- Development of a public conversational dataset ***MovieSent***
- A new conversational recommendation method ***ConvExtr***, which:
  - Estimates user's sentiment towards first 2 movies
  - Uses external dataset of reviews to predict user score towards the 3rd movie

MovieSent Dataset Construction

Based on the Coached Conversational Preference Elicitation dataset (CCPE)[1]

[1] "Coached Conversational Preference Elicitation" Radlinski et al. 2019

MovieSent Dataset Construction

1. Extracted conversations with at least 3 movies mentioned

MovieSent Dataset Construction

1. Extracted conversations with at least 3 movies mentioned
2. Movies labeled with RottenTomatoes IDs.

MovieSent Dataset Construction

1. Extracted conversations with at least 3 movies mentioned
2. Movies labeled with RottenTomatoes IDs.
3. Movies labeled with fine-grained user sentiment towards them:
   – Scale: [-3; +3] & None
   – 8 judges
   – 20% overlap

Examples of labeled utterances

Wizard utterance	User Utterance	Entity	Sentiment
What would be one of your favorite movies?	I love Mr. and Mrs. Smith. That's a great one.	mr_and_mrs_smith	3
Have you seen the Shape of Water?	I started watching that, but I just couldn't get into it enough to finish.	the_shape_of_water	-2
Have you seen Bridesmaids?	Nope.	bridesmaids_2011	None

MovieSent Dataset Statistics

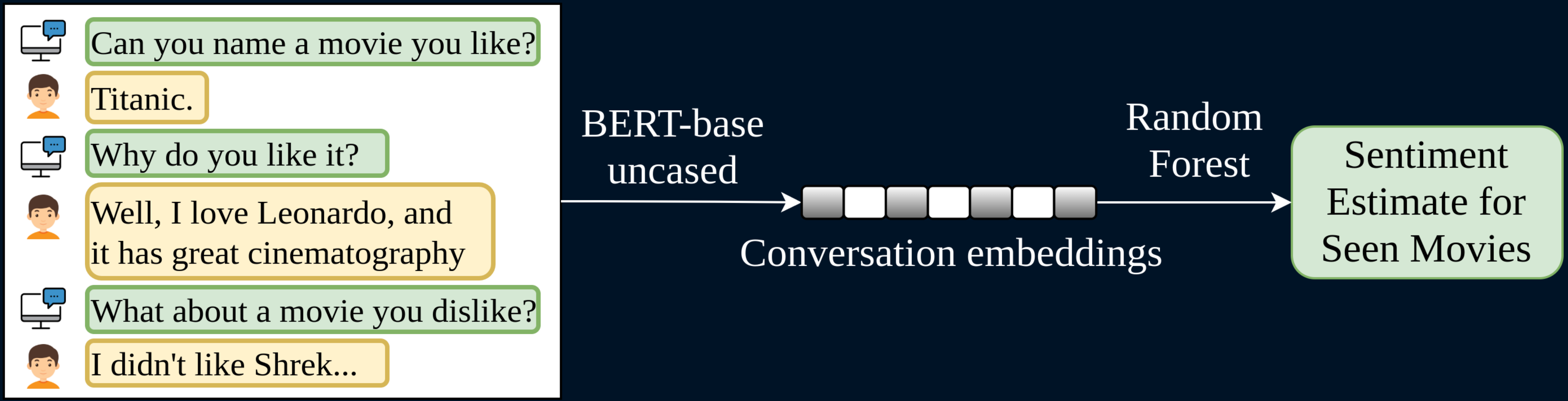
Conversations	489
Sentiment labels	2488
Unique entities	712
Weighted Cohen's κ	0.77

External Data

Scraped critics' reviews from RottenTomatoes:

Reviews	715,766
Critics	3,664
Median reviews	34
Unique Movies	42,423

Sentiment Estimation



Example Snippet Conversation from *MovieSent*

ConvExtr: General Idea

	Critic 1	Critic 2	$\cdots$	Critic k
Item 1	5	3	$\cdots$	
Item 2		2	$\cdots$	4
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Item T	3			1
Item S		4	$\cdots$	2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
Item N	4		$\cdots$	3

ConvExtr: General Idea

	Critic 1	Critic 2	$\cdots$	Critic k	Conv User
Item 1	5	3	$\cdots$		
Item 2		2	$\cdots$	4	
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	
Item T	3			1	~ 4.254
Item S		4	$\cdots$	2	~ 3.742
$\vdots$	$\vdots$	$\vdots$	$\ddots$		
Item N	4		$\cdots$	3	

ConvExtr: General Idea

	Critic 1	Critic 2	$\cdots$	Critic k	Conv User
Item 1	5	3	$\cdots$		
Item 2		2		4	
$\vdots$	$\vdots$			$\vdots$	
Item T	3			1	~ 4.254
Item S		1	$\cdots$	2	~ 3.742
$\vdots$	$\vdots$	$\vdots$	$\ddots$		
Item N	4		$\cdots$	3	

SVD++

ConvExtr: General Idea

	Critic 1	Critic 2	$\cdots$	Critic k	Conv User
Item 1	5	3	$\cdots$		
Item 2		2	$\cdots$	4	
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
Item T	3			1	~ 4.254
Item S			$\cdots$	2	~ 3.742
$\vdots$	$\vdots$	$\vdots$	$\ddots$		
Item N	4		$\cdots$	3	

SVD++

Prediction for Unseen movie

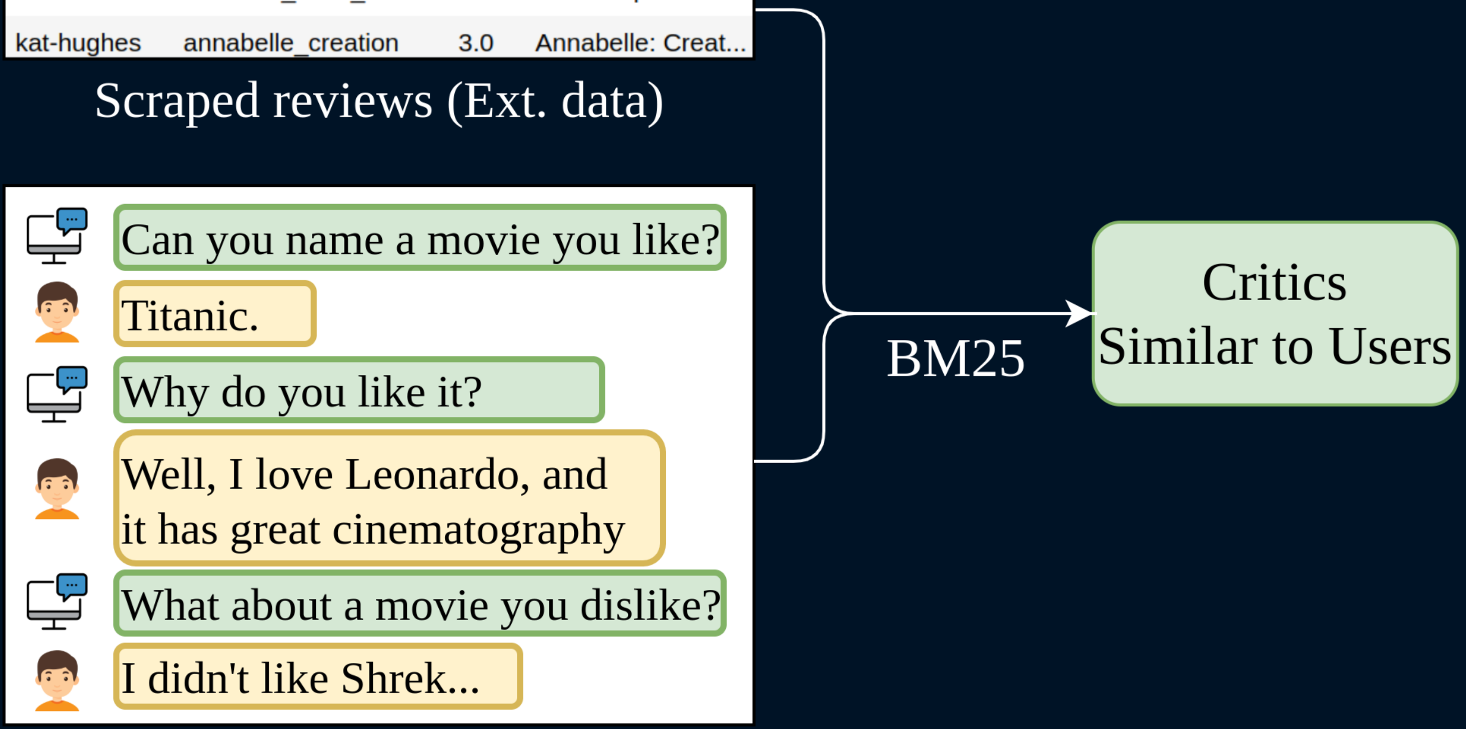
ConvExtr Model

Critics (Paid Professionals) $\neq$ Regular users

ConvExtr Model

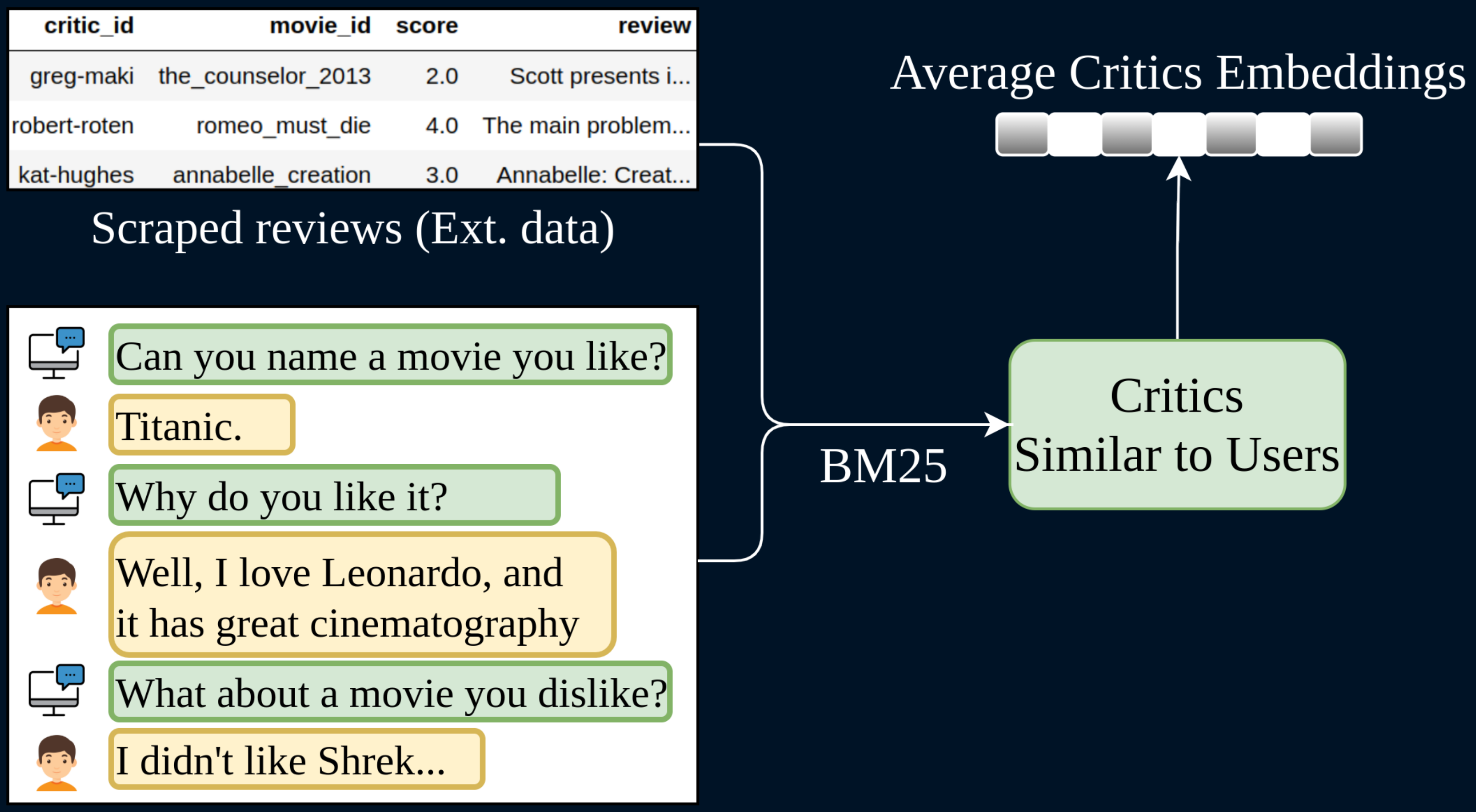
critic_id	movie_id	score	review
greg-maki	the_counselor_2013	2.0	Scott presents i...
robert-roten	romeo_must_die	4.0	The main problem...
kat-hughes	annabelle_creation	3.0	Annabelle: Creat...

Scraped reviews (Ext. data)



Example Snippet Conversation from *MovieSent*

ConvExtr Model

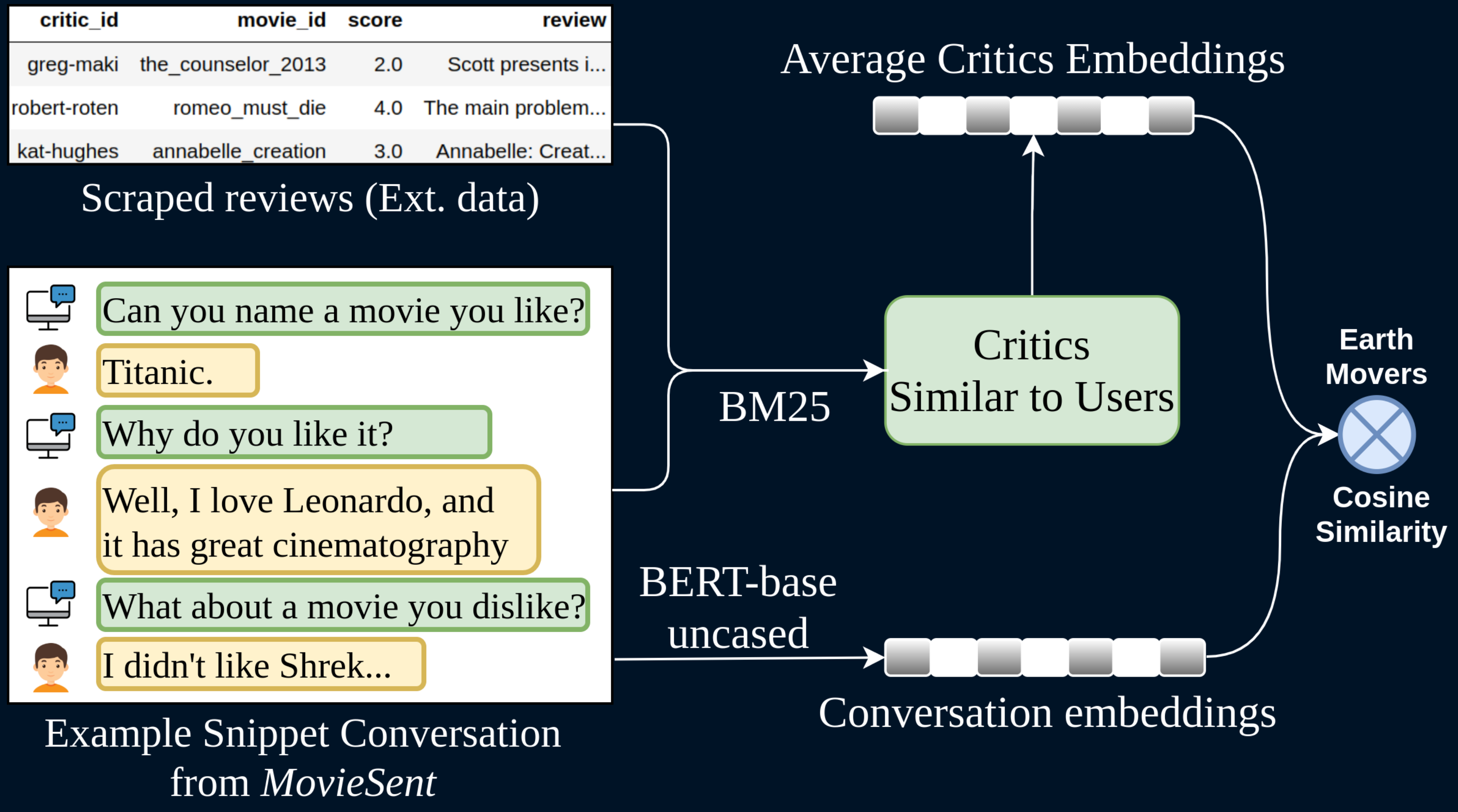


critic_id	movie_id	score	review
greg-maki	the_counselor_2013	2.0	Scott presents i...
robert-roten	romeo_must_die	4.0	The main problem...
kat-hughes	annabelle_creation	3.0	Annabelle: Creat...

Scraped reviews (Ext. data)

Example Snippet Conversation
from *MovieSent*

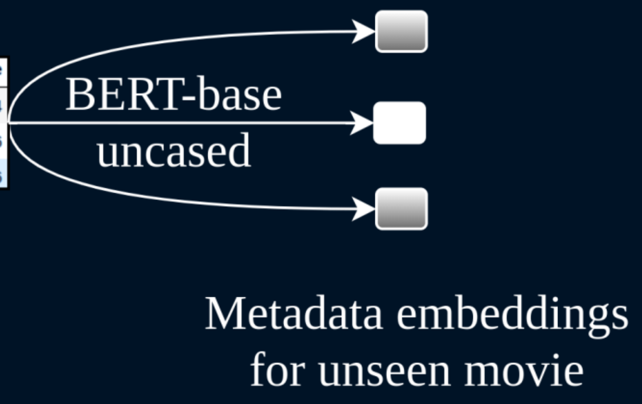
ConvExtr Model



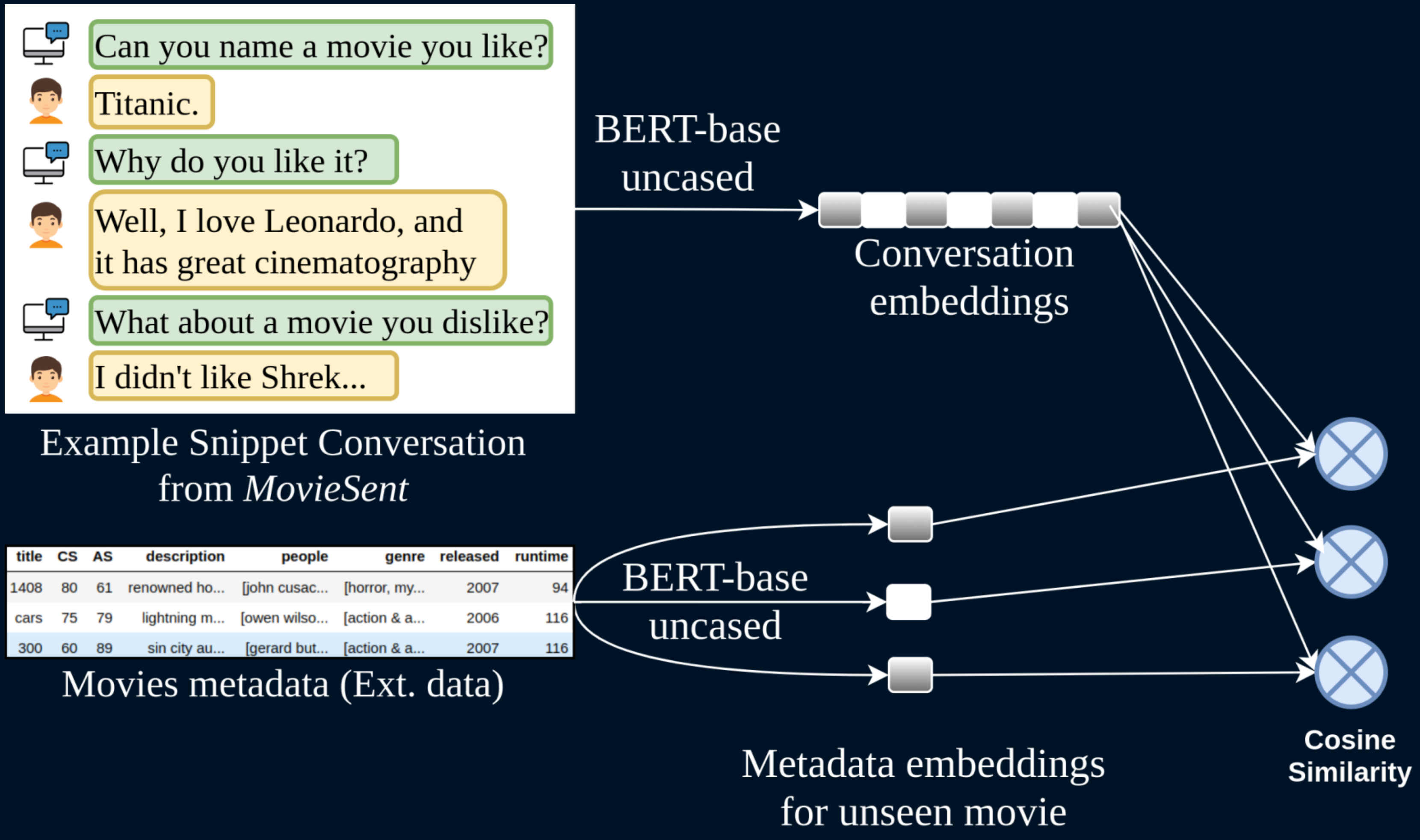
ConvExtr Model

title	CS	AS	description	people	genre	released	runtime
1408	80	61	renowned ho...	[john cusac...	[horror, my...	2007	94
cars	75	79	lightning m...	[owen wilso...	[action & a...	2006	116
300	60	89	sin city au...	[gerard but...	[action & a...	2007	116

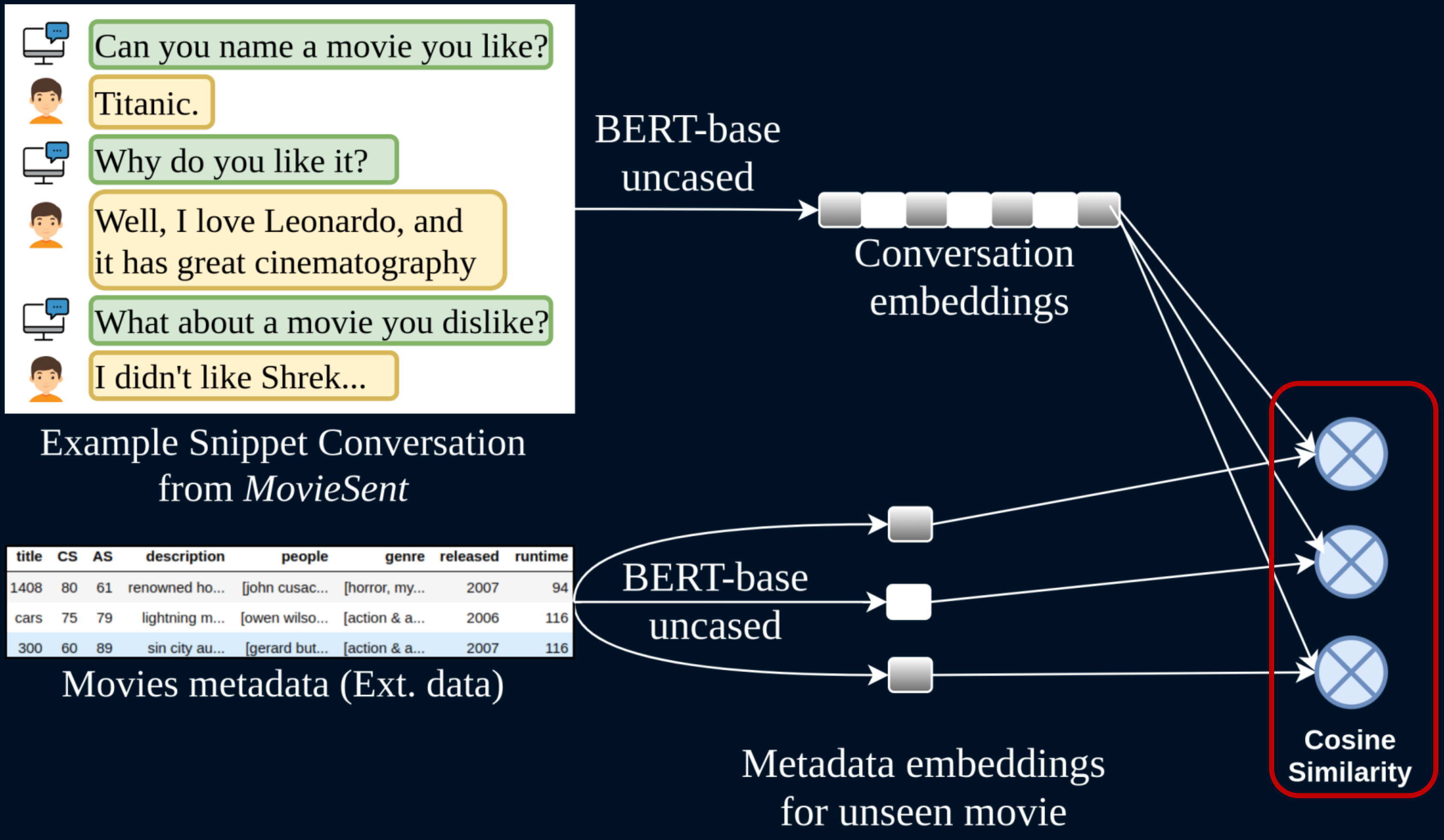
Movies metadata (Ext. data)



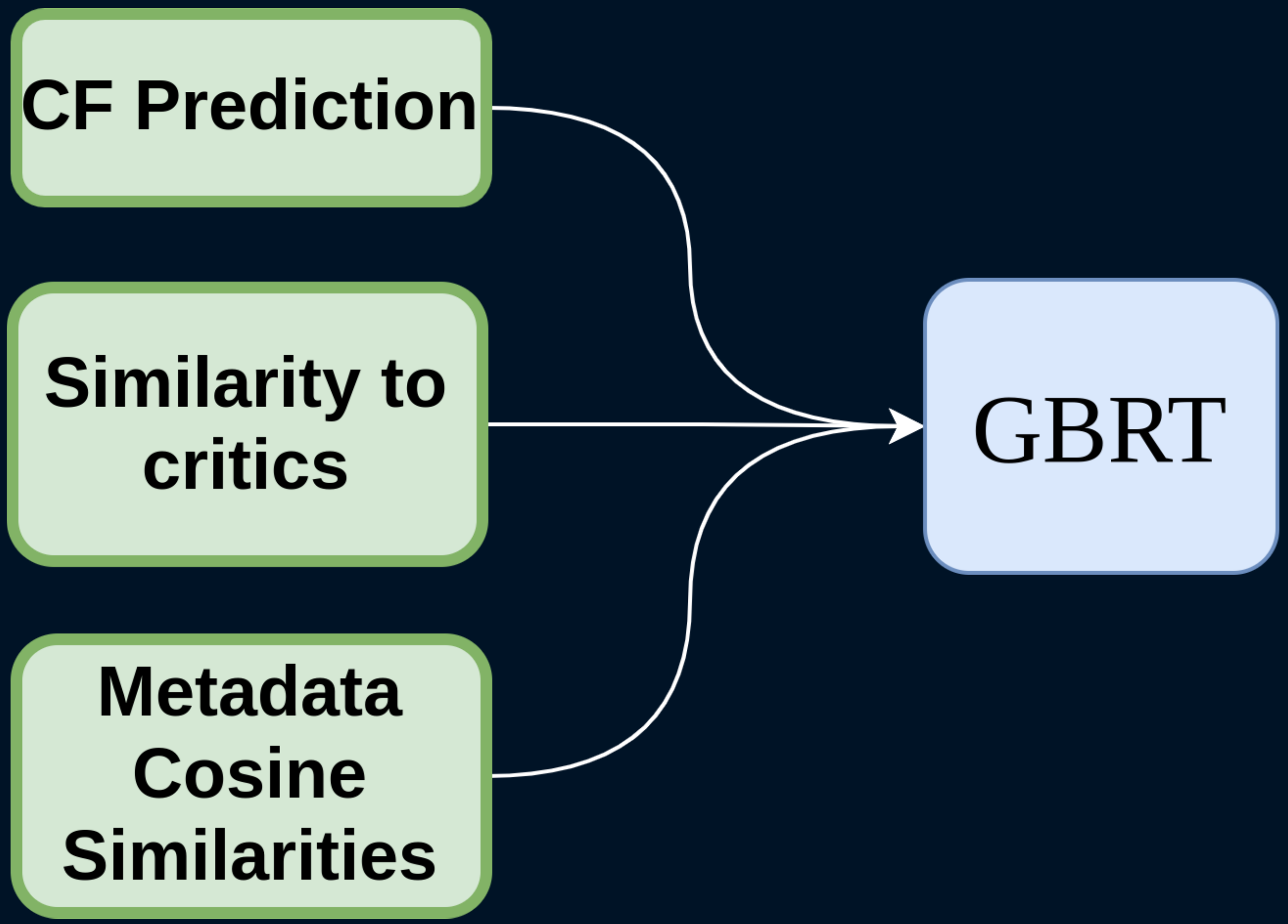
ConvExtr Model



ConvExtr Model



ConvExtr Model



Results

Model	RMSE	MAE
Baseline methods		
Average Critics	1.34	0.99
Average Audience	1.24	0.95

Results

Model	RMSE	MAE
Baseline methods		
Average Critics	1.34	0.99
Average Audience	1.24	0.95
ConvExtr (our method)		
KNN (CF only)	1.20	0.94
SVD (CF only)	1.18*	0.95
SVD++ (CF only)	1.14	0.92
GBRT	***1.09****	***0.84***

Results

Model	RMSE	MAE
Baseline methods		
Average Critics	1.34	0.99
Average Audience	1.24	0.95
ConvExtr (our method)		
KNN (CF only)	1.20	0.94
SVD (CF only)	1.18*	0.95
SVD++ (CF only)	1.14	0.92
GBRT	***1.09****	***0.84***
Best Possible:	0.84	0.64

Conclusion

- Using conversation to select more similar users for CF improves recommendation performance
- The resulting insights offer a promising direction for improving conversational recommendation systems

- Dataset and code available at:

Acknowledgements

This work was partially supported by a grant from Amazon Alexa towards the study of conversational search and recommendation.

amazon alexa

Thank you