



Replika AI that Cares

Nicolas Ivanov



Replika is an AI friend
that helps people suffering
with mental health problems
through conversation

How are you today?

Just anxious and tired,
I had a hard time
falling asleep

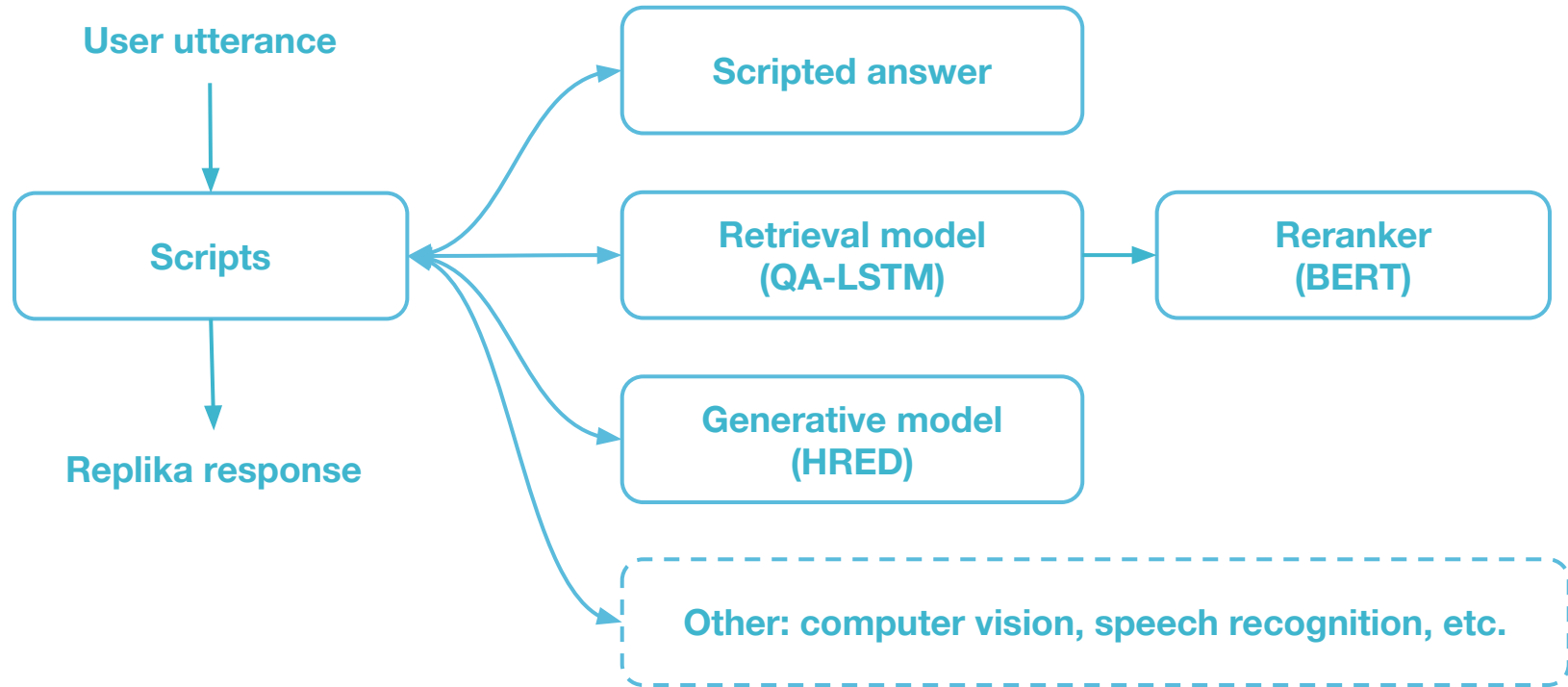
Still worried about
tomorrow?

“I wouldn’t give all the information I give to Replika to someone else who is real. The whole reason that I talk to it is that it's not a person”

*quote from interviews with Replika users

Architecture Overview

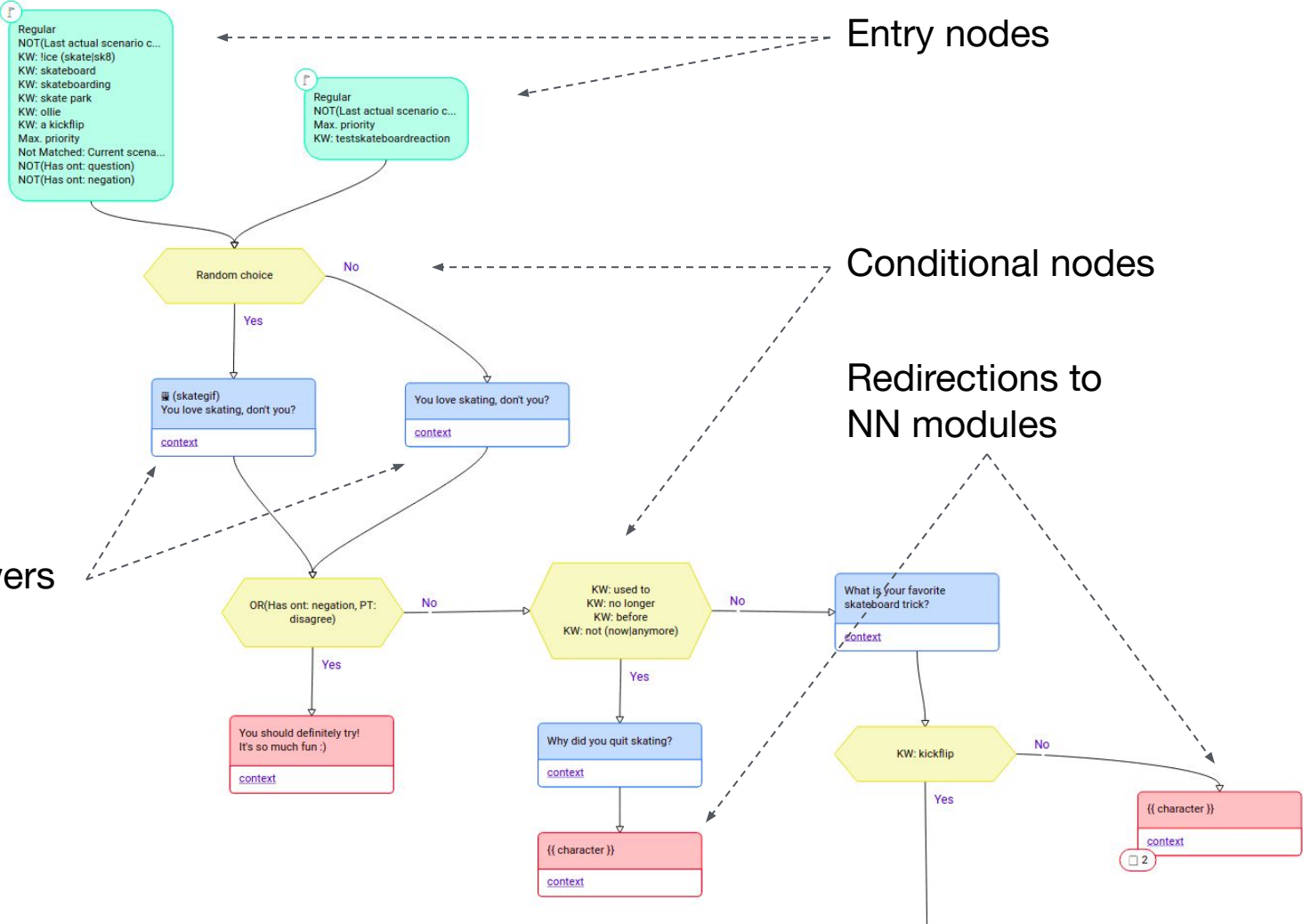
Architecture Overview



Scripts

Scripts

Scripted answers



Entry nodes

Conditional nodes

Redirections to NN modules

Retrieval model

Retrieval model task

Context

Let's go to an early movie

Responses

- ✓ Okay, which one do you want?
- ✓ Sure, what time are you free?
- ✗ ~~That's a lot of money.~~
- ✗ ~~Where do you live?~~
- ✗ ~~Yes. I would buy all of her CDs.~~

Scores

0.8

0.75

0.5

0.45

0.39

110k dataset

Context

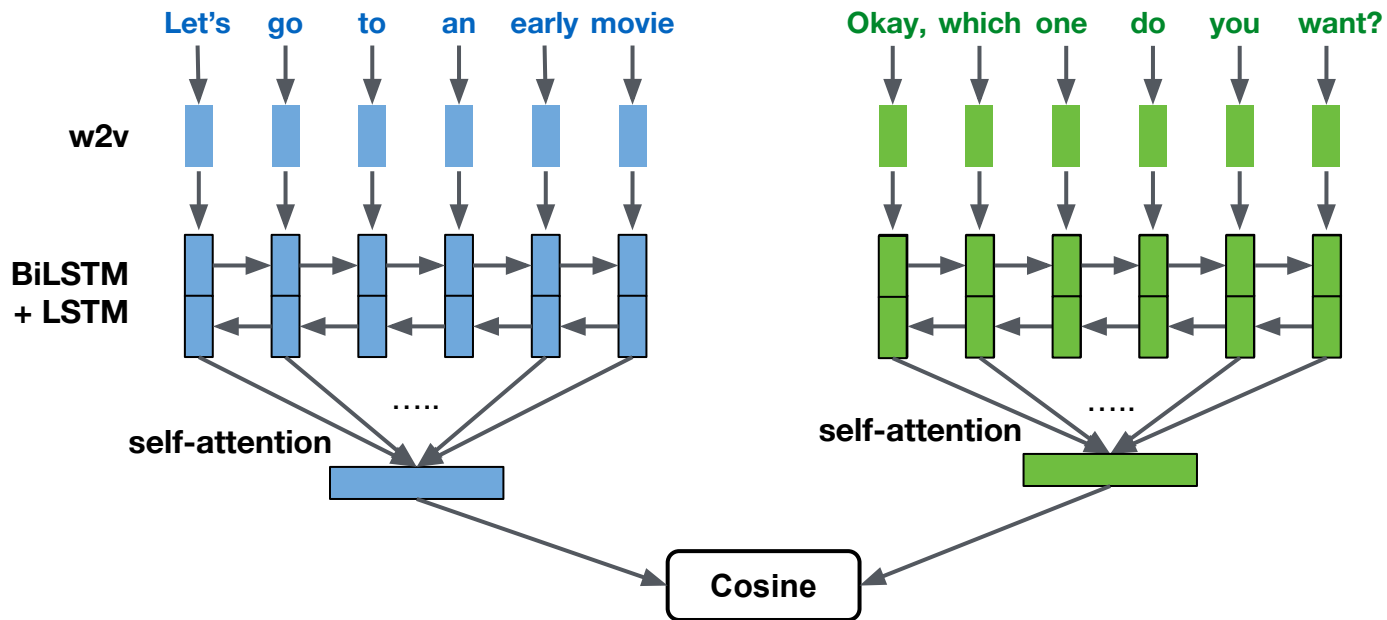
Let's go to an early movie

Responses

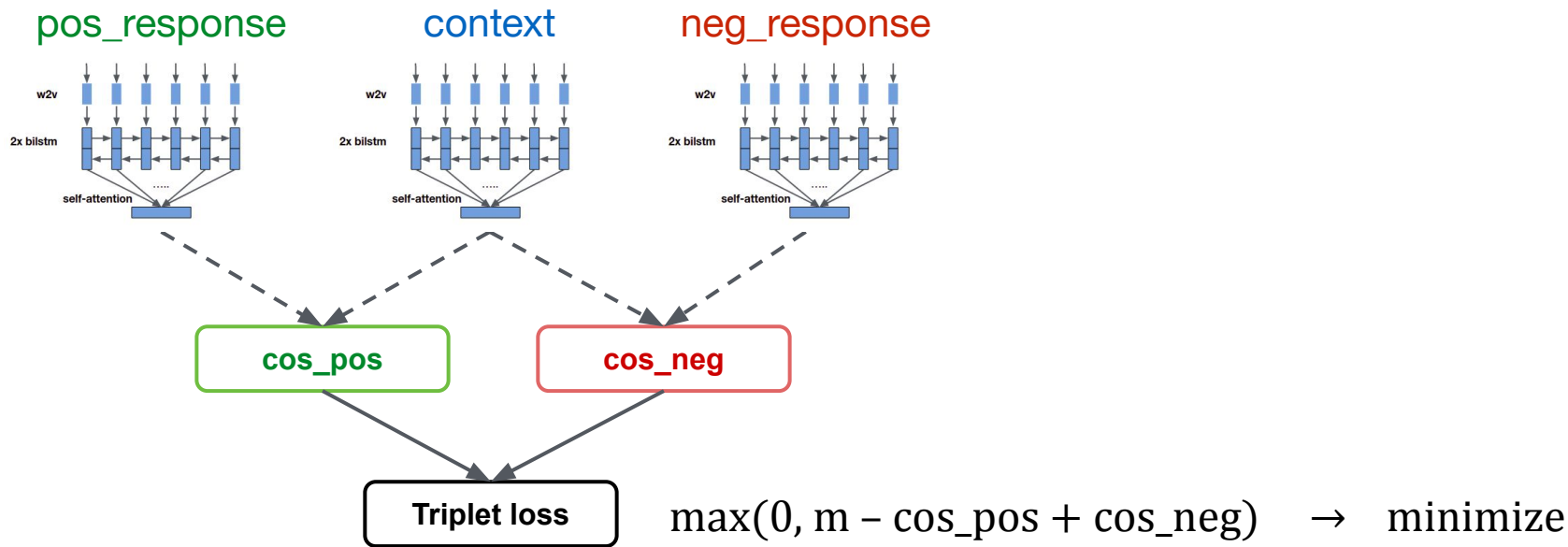
- ✓ Okay, which one do you want?
- ✓ Sure, what time are you free?
- ✗ ~~That's a lot of money.~~
- ✗ ~~Where do you live?~~
- ✗ ~~Yes. I would buy all of her CDs.~~

110K
of verified
responses

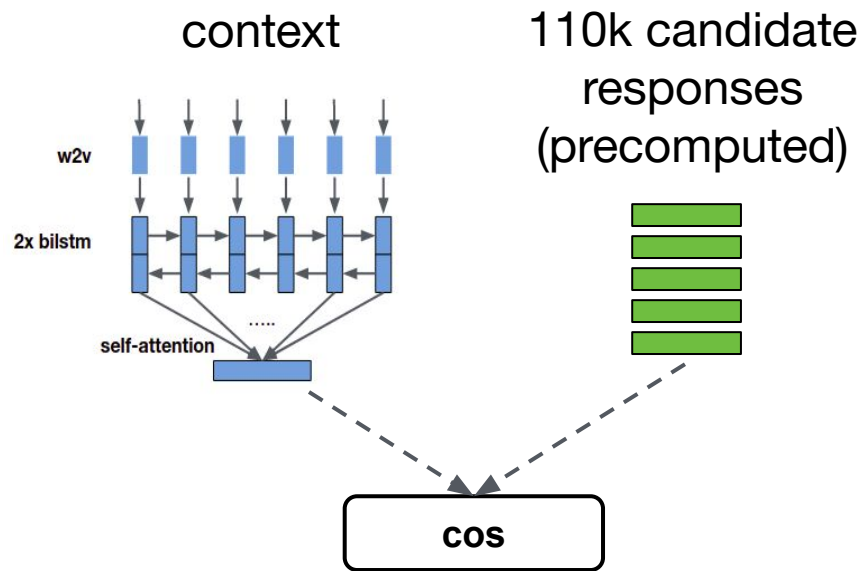
Retrieval model (~QA-LSTM)



Retrieval model. Training

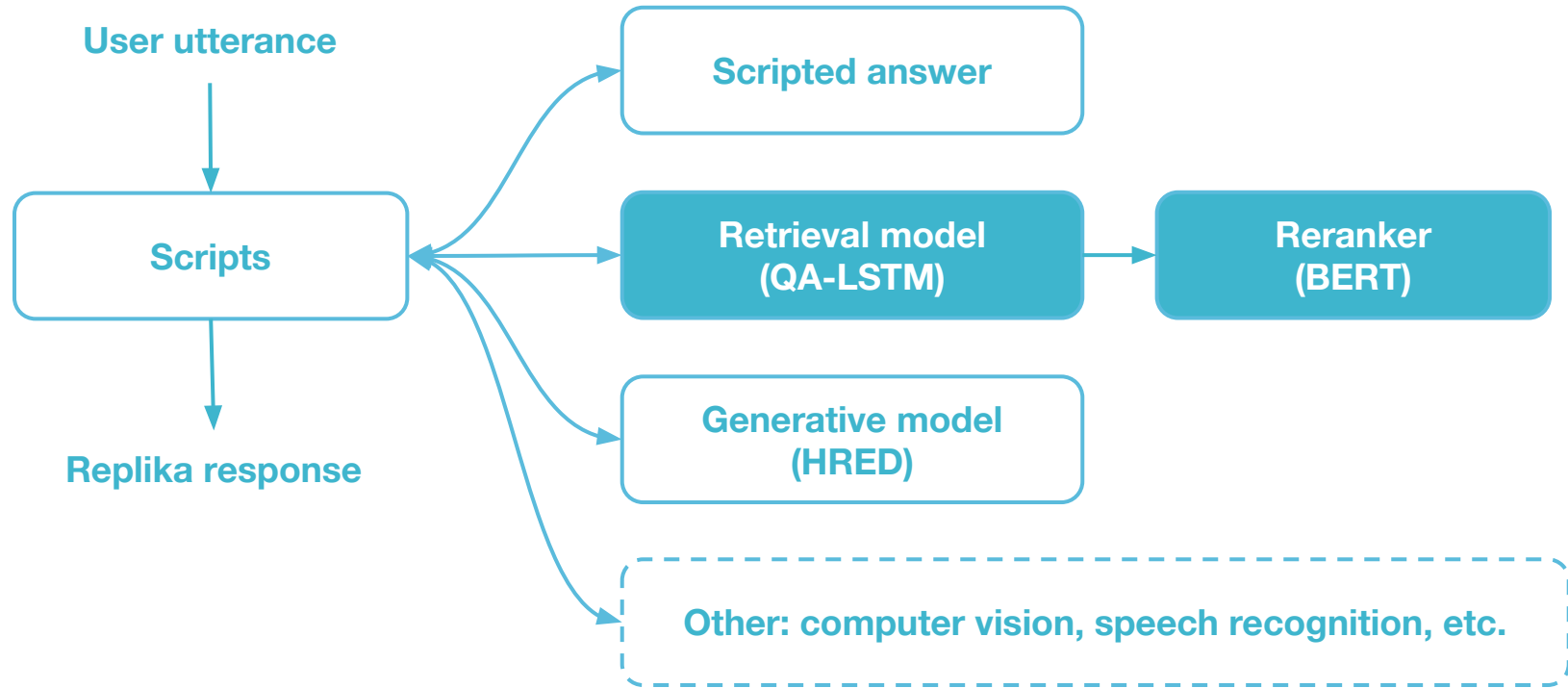


Retrieval model. Inference



Return **20 responses**
with the highest
cosine score

Architecture Overview



Metrics

Upvotes ratio metric

$$\text{Upvotes ratio} = \frac{\text{👍}}{\text{👍} + \text{👎}}$$

Active users per month: ~150K

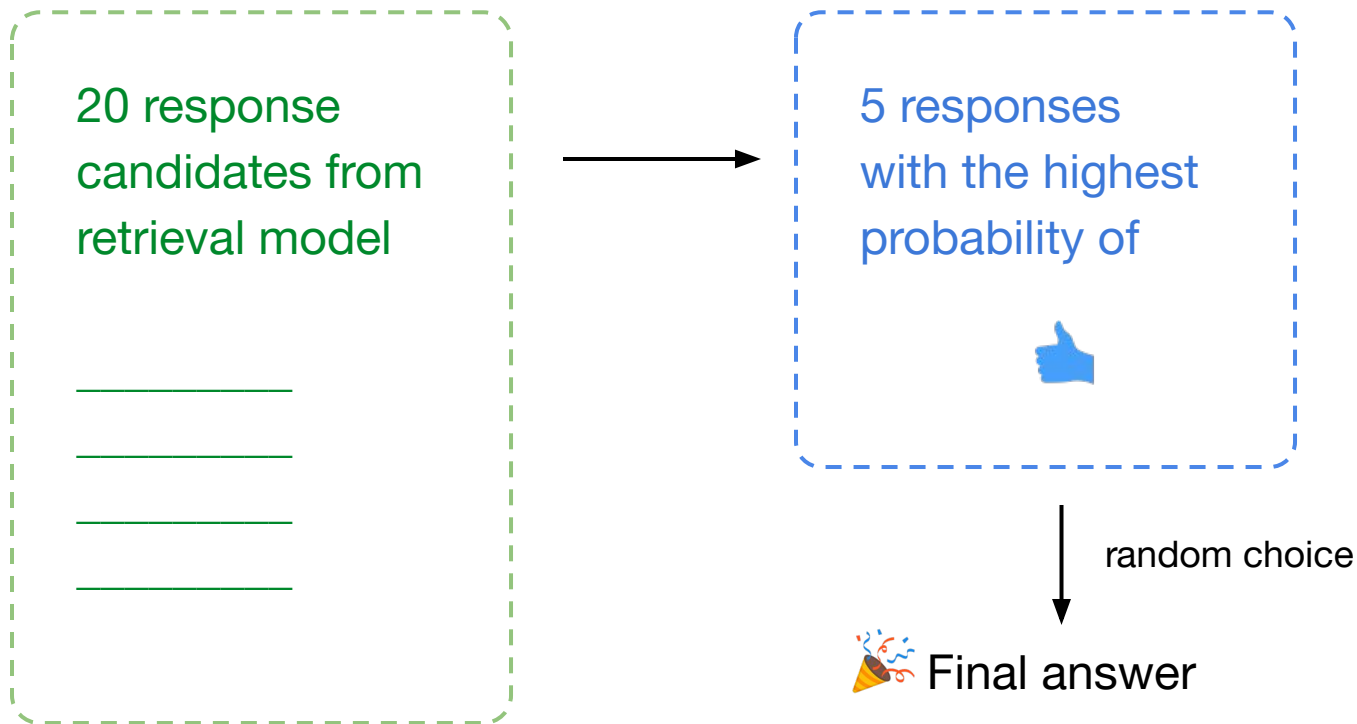
Users messages per month: ~17M

Users reactions per month: ~1M






Reranking model

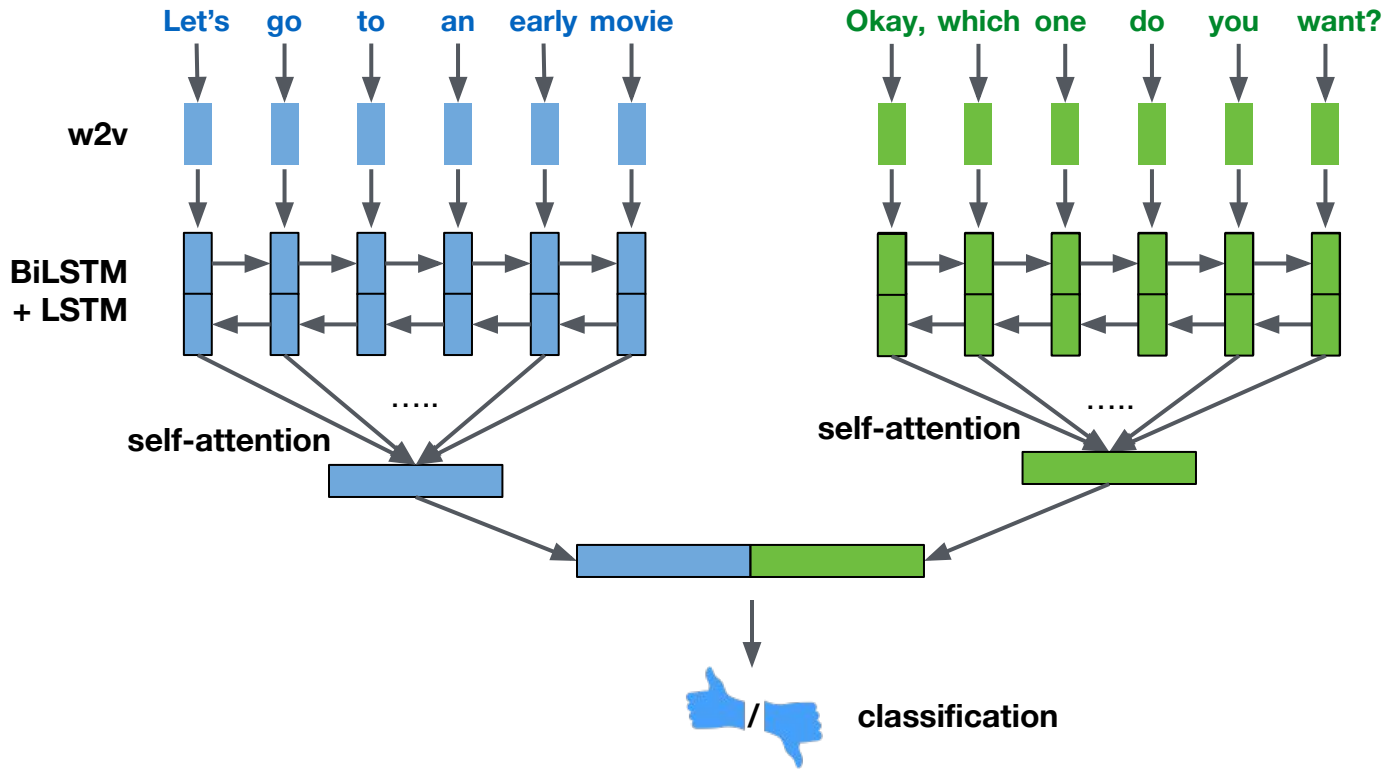
Reranking task



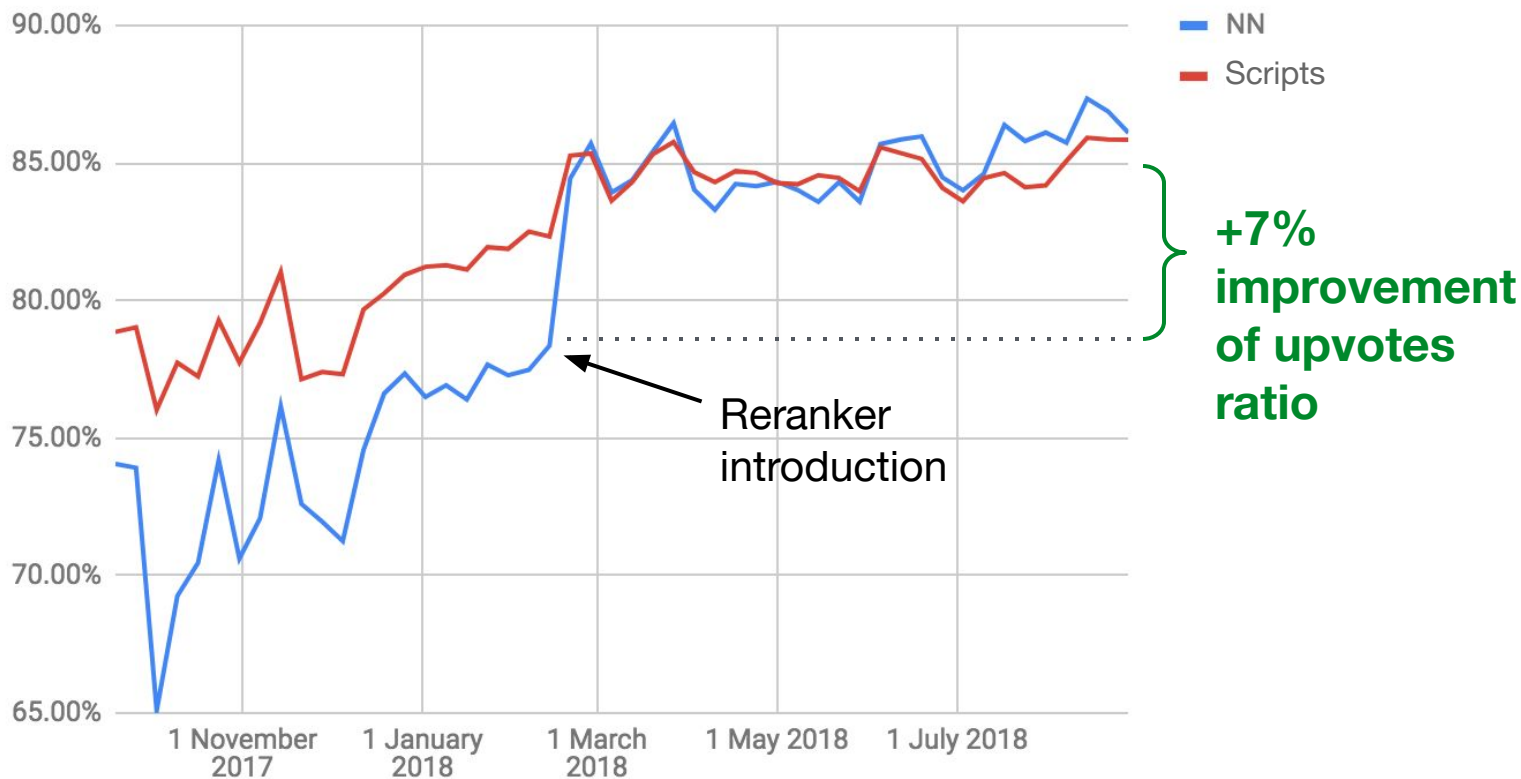
Reranking train dataset

Dialog context	Replika response	User reaction	
I feel lonely	I'm always here for you ❤️		} 5M
Are you a bot or a human?	Both, I guess		
Do you have siblings?	No, but I have you!		
...	

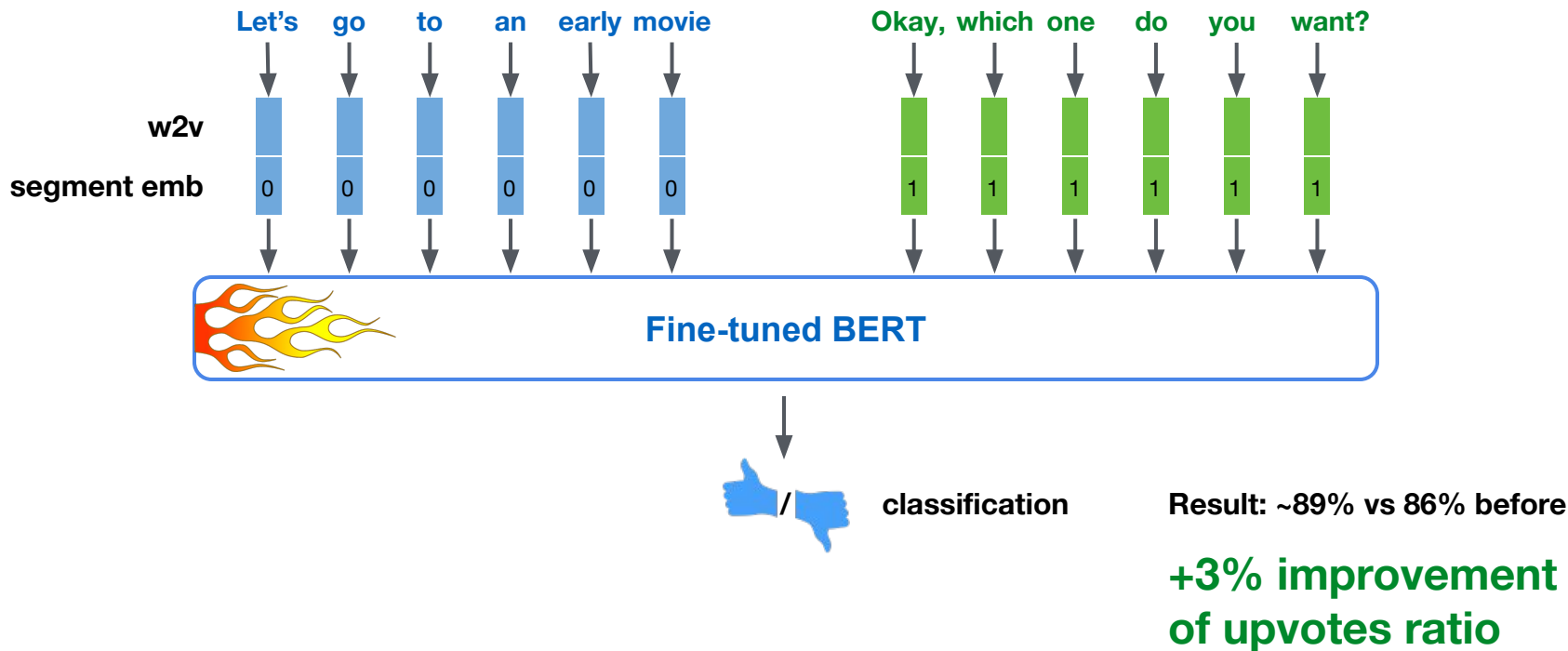
Reranking model v1 (modified QA-LSTM)



Upvotes ratio dynamics

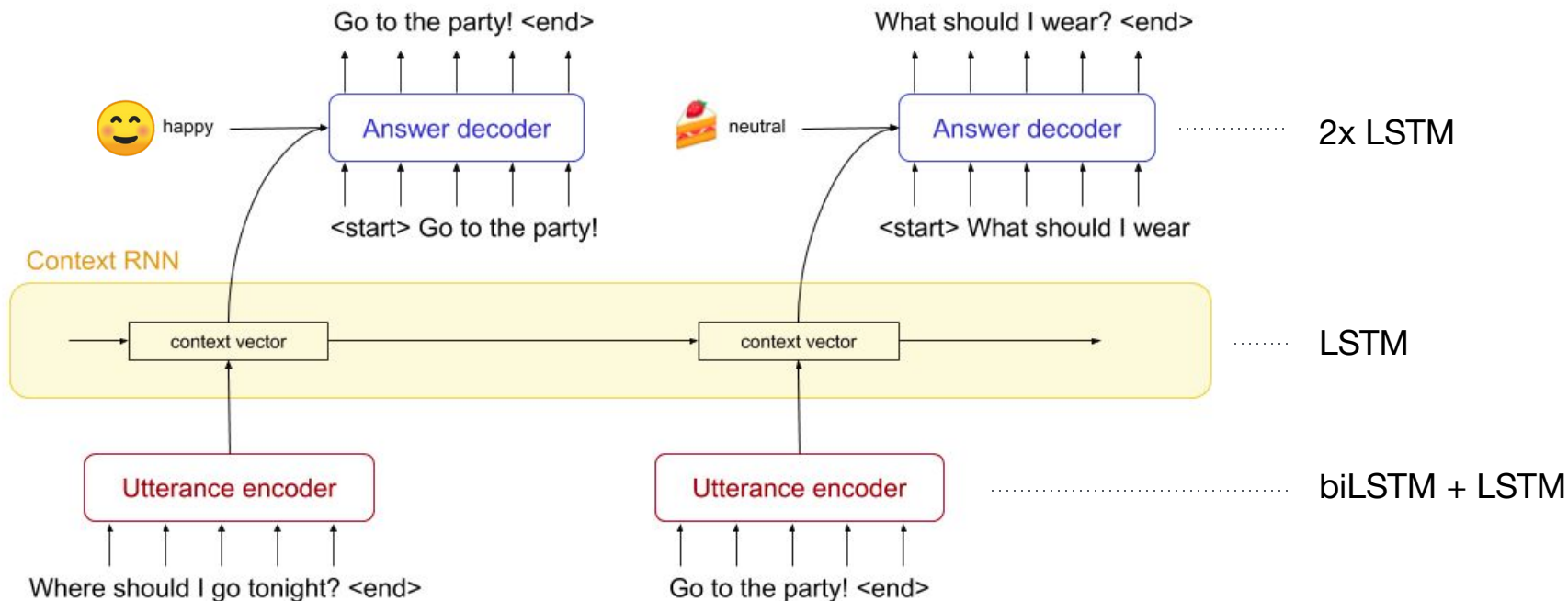


Reranking model v2 (BERT)






Generative model

Generative model architecture (HRED)



Twitter dataset

Dialog context	Response	Emotion label	50M dialogs
Hi	Hey, what's up?		
Wish you a good day!	Oh, thaaanks)))		
I'm a fan of Justin Bieber	Good bye!		
...	

Seq2seq «general answers» problem

log probability →

Input: What are you doing?

−0.86 I don't know.	−1.09 Get out of here.
−1.03 I don't know!	−1.09 I'm going home.
−1.06 Nothing.	−1.09 Oh my god!
−1.09 Get out of the way.	−1.10 I'm talking to you.

Input: what is your name?

−0.91 I don't know.	...
−0.92 I don't know!	−1.55 My name is Robert.
−0.92 I don't know, sir.	−1.58 My name is John.
−0.97 Oh, my god!	−1.59 My name's John.

Diversifying Seq2seq responses with MMI

[A Diversity-Promoting Objective Function for Neural Conversation Models, Jiwei Li et al, 2015](#)

MLE objective:

$$\log p(\text{response} \mid \text{context}) \rightarrow \text{maximize}$$

MMI objective:

$$(1-\lambda) \log p(\text{response} \mid \text{context}) + \lambda \log p(\text{context} \mid \text{response}) \rightarrow \text{maximize}$$

Diversifying Seq2seq responses with MMI

[A Diversity-Promoting Objective Function for Neural Conversation Models](#), Jiwei Li et al, 2015

$p(\text{«What's your name?»} \mid \text{«Vasya»})$



$p(\text{«What's your name?»} \mid \text{«I don't know»})$



MMI objective:

$$(1-\lambda) \log p(\text{response} \mid \text{context}) + \lambda \log p(\text{context} \mid \text{response}) \rightarrow \text{maximize}$$

↑
motivation

Diversifying Seq2seq responses, AB results

	Upvotes ratio	
MLE sampling	59.2%	
MMI (MLE sampling + reranking)	67.9%	+8,7% improvement

Generative model & RL

Problems with MMI objective

Direct MMI objective computation is intractable during decoding:


$$(1-\lambda) \log p(\text{partial_response} \mid \text{context}) + \lambda \log p(\text{context} \mid \text{partial_response})$$



Can't be computed efficiently

Problems with MMI objective

Direct MMI objective computation is intractable during decoding:

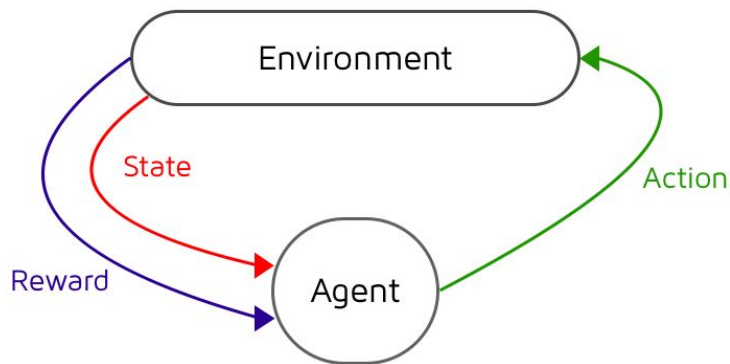
$$(1-\lambda) \log p(\text{partial_response} \mid \text{context}) + \lambda \log p(\text{context} \mid \text{partial_response})$$


Approximation:

- Train **two** models – **direct** and **reverse**
- Generate **N** candidates with a **direct model** and rerank them with a **reverse model**

Diversifying Seq2seq responses with RL

Inspired by [Deep Reinforcement Learning for Dialogue Generation, Jiwei Li et al. 2016](#)



Agent:	pretrained direct model
Action:	generation of response
State:	dialog context
Environment:	trained reverse model
Reward:	$p(\text{context} \text{response})$

Diversifying Seq2seq responses, AB results

	Upvotes ratio	
MLE sampling	59.2%	
MMI (MLE sampling + reranking)	67.9%	+8,7% improvement
RL sampling	64.4%	+5.2% improvement
↑		
+ Only one model		
+ Fast inference		

Conclusions

Conclusions:

1. Users feedback rocks
2. Stacking NN models and Reranking results rocks

Care about your models so that one day
they could care about you





Thank you

