The Conversational Intelligence Challenge 2 (ConvAI2)

Transfer-Transfo: A Transfer Learning Approach for Neural Network Conversational Agents

thomas@huggingface.co

The Conversational Intelligence Challenge 2 (NeurIPS 2018 competition)

Persona 1	Persona 2
I like to ski	I am an artist
My wife does not like me anymore	I have four children
I have went to Mexico 4 times this year	I recently got a cat
I hate Mexican food	I enjoy walking for exercise
I like to eat cheetos	I love watching Game of Thrones

[PERSON 1:] Hi
[PERSON 2:] Hello ! How are you today ?
[PERSON 1:] I am good thank you , how are you.
[PERSON 2:] Great, thanks ! My children and I were just about to watch Game of Thrones.
[PERSON 1:] Nice ! How old are your children?
[PERSON 2:] I have four that range in age from 10 to 21. You?
[PERSON 1:] I do not have children at the moment.
[PERSON 2:] That just means you get to keep all the popcorn for yourself.
[PERSON 1:] And Cheetos at the moment!
[PERSON 2:] Good choice. Do you watch Game of Thrones?
[PERSON 1:] No, I do not have much time for TV.
[PERSON 2:] I usually spend my time painting: but, I love the show.

Example dialog from the PERSONA-CHAT dataset. Person 1 is given their own persona (top left) at the beginning of the chat, but does not know the persona of Person 2, and vice-versa. They have to get to know each other during the conversation.

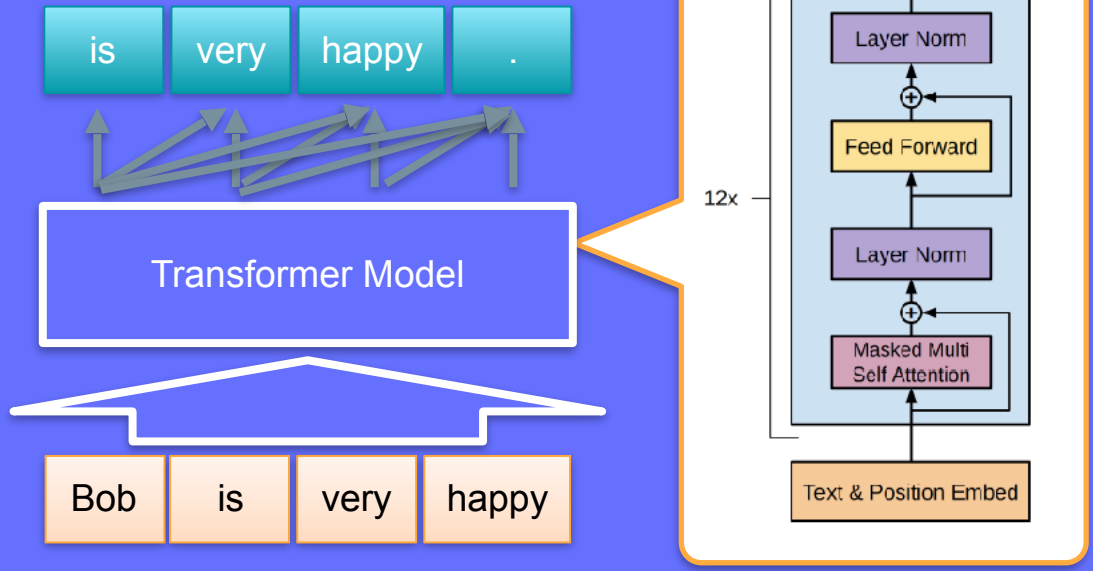
Chit-Chatting with a coherent personality

Open domain dialog
Short conversation: <10 turns
Small talk: shallow topics, quick switches

A Generative Transformer Model for Chit-Chat 🤖

- Input and outputs: Fixed-length sequences of tokens (« words », in our case BPE)
- Each output is a probability distribution for the next token in the sequence over the vocabulary of tokens.

- Inside:
  - Self-attention layers: compute Key-Value-Query for each element of the sequence
  - Decoder Transformer: Causal masking in the attention heads to only attend to the past.



Training a Neural Network Model for Chit-Chat

- PERSONA-CHAT is **one of the biggest** multi-turn dialog dataset :
  - 164,356 utterances and about 1-2M words
  - Average number of turns: 14
- But it is still **small** for training a deep learning model:
  - 1B words in the Billion Words dataset
  - ~1M sentences in CoNLL 2012 (used for training co-reference systems)
- And generating an engaging open-domain dialogue requires:
  - topic-coherence,
  - dialogue-flow,
  - common-sense,
  - short term memory,
  - co-reference resolution,
  - sentimental analysis,
  - textual entailment…

What can we do? 

Transfer Learning 

Step 1: Pre-training – Learn Natural Language Generation 🦄

Pre-train the model with a **language modeling objective** on a **large dataset** of contiguous span of texts.

- This 1st stage of training learns the initial parameters of the neural network model.
- The combination of the language modeling objective and un-shuffled dataset was shown to provide the model with some kind of world knowledge and an ability to process long-range dependencies.
- More recently these kind of pre-trained Transformer models have been shown to improve the state of the art on many downstream NLP tasks and in particular in commonsense-knowledge tasks.
- In our experiments, we started from the model of *Radford et al.* pre-trained on the Toronto Book dataset with ~7k books.

A Simple Method for Commonsense Reasoning by Trinh & Le (2018), *Improving, Language Understanding by Generative Pre-Training* by Radford et al. (2018), *Universal Language Model Fine-tuning for Text Classification* by Howard and Ruder (2018), *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* by Jacob Devlin et al (2018)

Step 2: Fine-tuning – Learn How to Hold a Conversation 

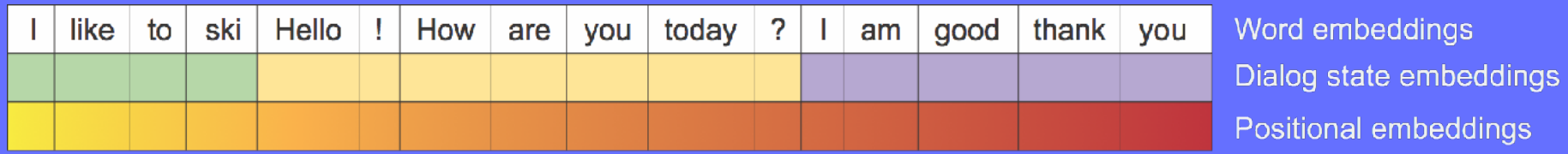
Encoding a Dialog and a Persona

- Now that our model is provided with basic common-sense and some capabilities to generate coherent long-range linguistic structures, we need to teach it the specificities of dialog:
  - Alternating utterances
  - Dialog flow (« speech/dialog acts »)
  - Conditioning on a personality
- How to build a sequential inputs for our model from a conditioned dialog?
  - Unlike RNN/LSTM, Transformers don’t possess an inherent notion of sequentiality and position
  - We need to add positional embeddings to incorporate sequentiality

Encoding a Dialog and a Persona

- Now that we have a model with basic common-sense and co-reference capabilities, we need to teach it the specificities of dialog:
  - Alternating utterances
  - Dialog flow (« speech/dialog acts »)
  - Conditioning on a personality

- How to build a sequential inputs for our model from a conditioned dialog?
  - Unlike RNN/LSTM, Transformers don't possess a natural notion of sequentiality and position
  - We need to add positional embeddings to incorporate sequentiality
  - We add special embeddings related to utterances and personas



Encoding a Dialog and a Persona

- We can play with these embeddings to manipulate the notion of a sequence

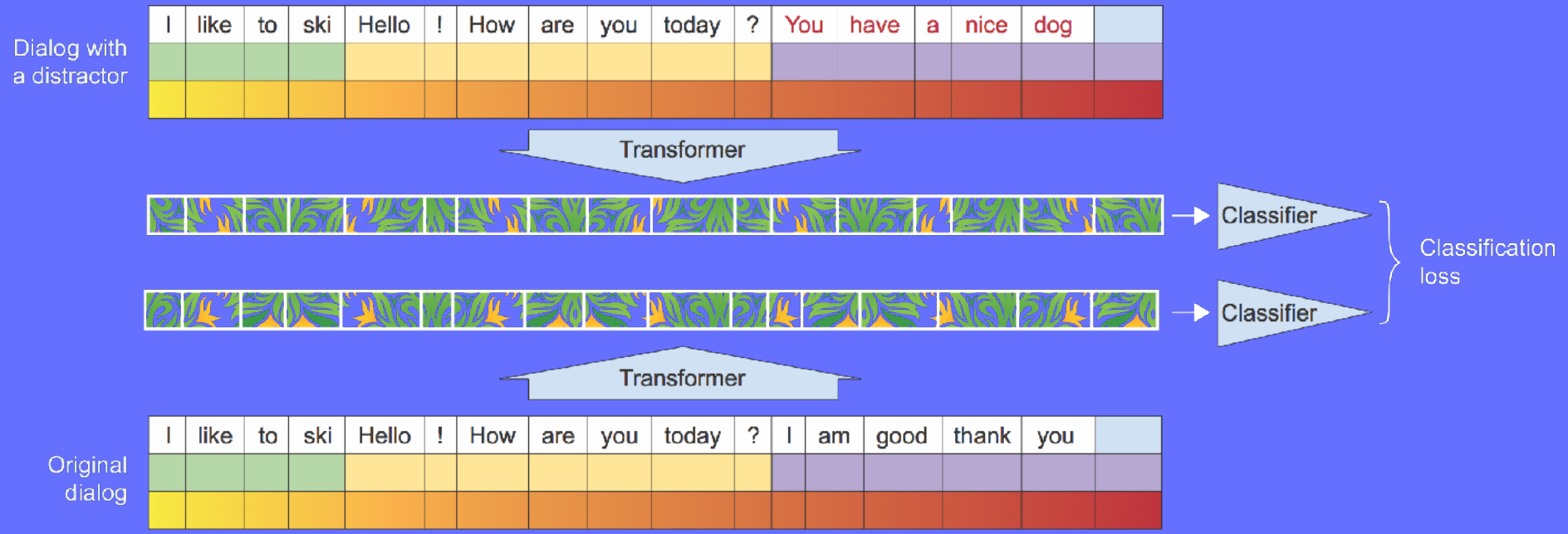
Repeating specific embeddings to control positioning information

- We can also augment the dataset to bias towards positional invariance

Permutation augmented dataset to bias towards positional invariance

Semantic Learning on Dialog Utterances

- Learning to distinguish a real answer from a distractor.



Combined with language modeling fine-tuning in a multi-task fashion

Results

Validation set (public) Leaderboard – Test set (hidden) Leaderboard

Model	Creator	PPL	Hits@1	F1
	(Hugging Face)	17.51	82.1	19.09
	Happy Minions	29.85	-	17.79
	ADAPT Centre	22.57	-	20.3
	Lost in Conversation	-	17.3	17.79
	Khai Mai Alt	-	85.3	17.64
	Pinta	23.86	-	17.27
	Mohd Shadab Alam	34.12	13.4	17.08
	Sonic	38.87	-	16.88
	NEUROBOTICS	39.7	-	16.82
	1st-contact	36.54	13.3	16.58
topicSeq2seq	Team Pat	-	-	16.58
	Roboy	-	-	16.25
	Tensorborne	44.64	12.1	16.13
	flooders	-	-	15.96
	Clova Xiaodong Gu	-	-	15.39
	IamNotAdele	53.46	-	12.85
	Little Baby(AI小奶娃)	-	83.0	-
	High Five	-	79.1	-
	Sweet Fish	-	75.6	-
	Cats'team	-	43.4	-
	loopAI	-	29.7	-
	Salty Fish	33.46	-	-

Rank	Creator	PPL	Hits@1	F1
1	(Hugging Face)	16.28	80.7	19.5
2	ADAPT Centre	31.4	-	18.39
3	Happy Minions	29.01	-	16.01
4	High Five	-	65.9	-
5	Mohd Shadab Alam	29.94	13.8	16.91
6	Lost in Conversation	-	17.1	17.77
7	Little Baby(AI小奶娃)	-	64.8	-
8	Sweet Fish	-	45.7	-
9	1st-contact	31.98	13.2	16.42
10	NEUROBOTICS	35.47	-	16.68
11	Cats'team	-	35.9	-
12	Sonic	33.46	-	16.67
13	Pinta	32.49	-	16.39
14	Khai Mai Alt	-	34.6	13.03
15	loopAI	-	25.6	-
16	Salty Fish	34.32	-	-
17	Team Pat	-	-	16.11
18	Tensorborne	38.24	12.0	15.94
19	Team Dialog 6	40.35	10.9	7.27
20	Roboy	-	-	15.83
21	IamNotAdele	66.47	-	13.09

That’s it for today
Thanks for listening!

thomas@huggingface.co