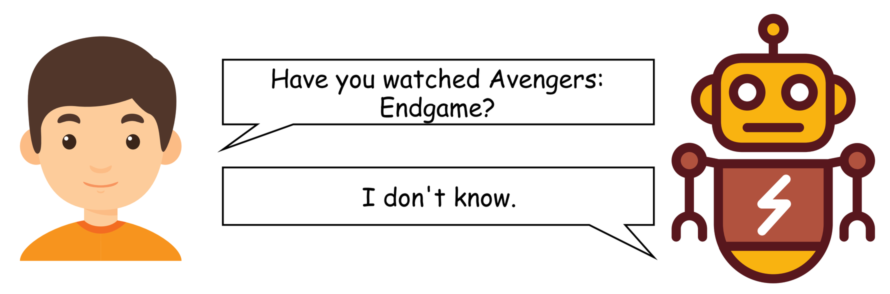
Improving Neural Response Diversity with Frequency-Aware Cross-Entropy Loss

Shaojie Jiang Pengjie Ren Christof Monz Maarten de Rijke

University of Amsterdam

Summary

Task: We address the low-diversity problem of Sequence-to-Sequence (Seq2Seq) based chatbots – the problem of frequently generating dull responses like "I don't know" or "I'm sorry".



Method: Other than previous diagnoses, we believe this problem is also caused by frequent tokens. Therefore, we propose a new loss function to counter the effect of different token frequencies.

Main contributions:

- We show that frequency variance of different tokens can cause model over-confidence and low response diversity.
- We propose a Frequency-Aware Cross-Entropy (FACE) loss function to balance per-token training loss, which alleviates model over-confidence and, hence, improves response diversity.
- We investigate two token frequency calculation methods and corresponding frequency-based weighting mechanisms for FACE.

Problem Analysis

Model over-confidence and low-diversity are statistical and empirical symptoms of the same problem: imbalance of training losses caused by token frequency variances.

What is model over-confidence?

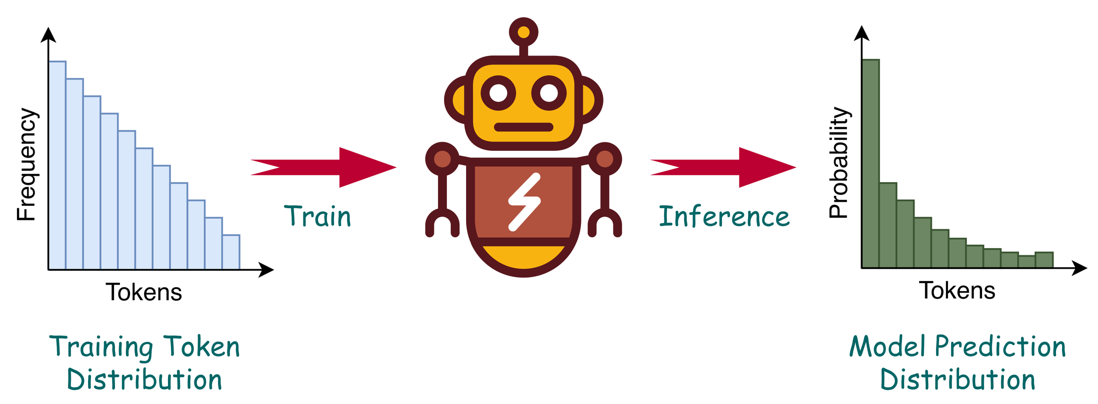


Figure: Token frequency and model prediction

What is the result of model over-confidence?

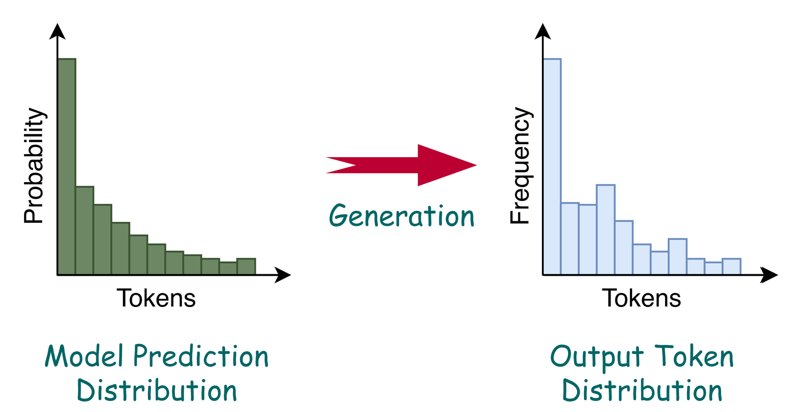


Figure: Predicted probability and model generation

Frequent tokens in training data will result in frequent tokens in model outputs.

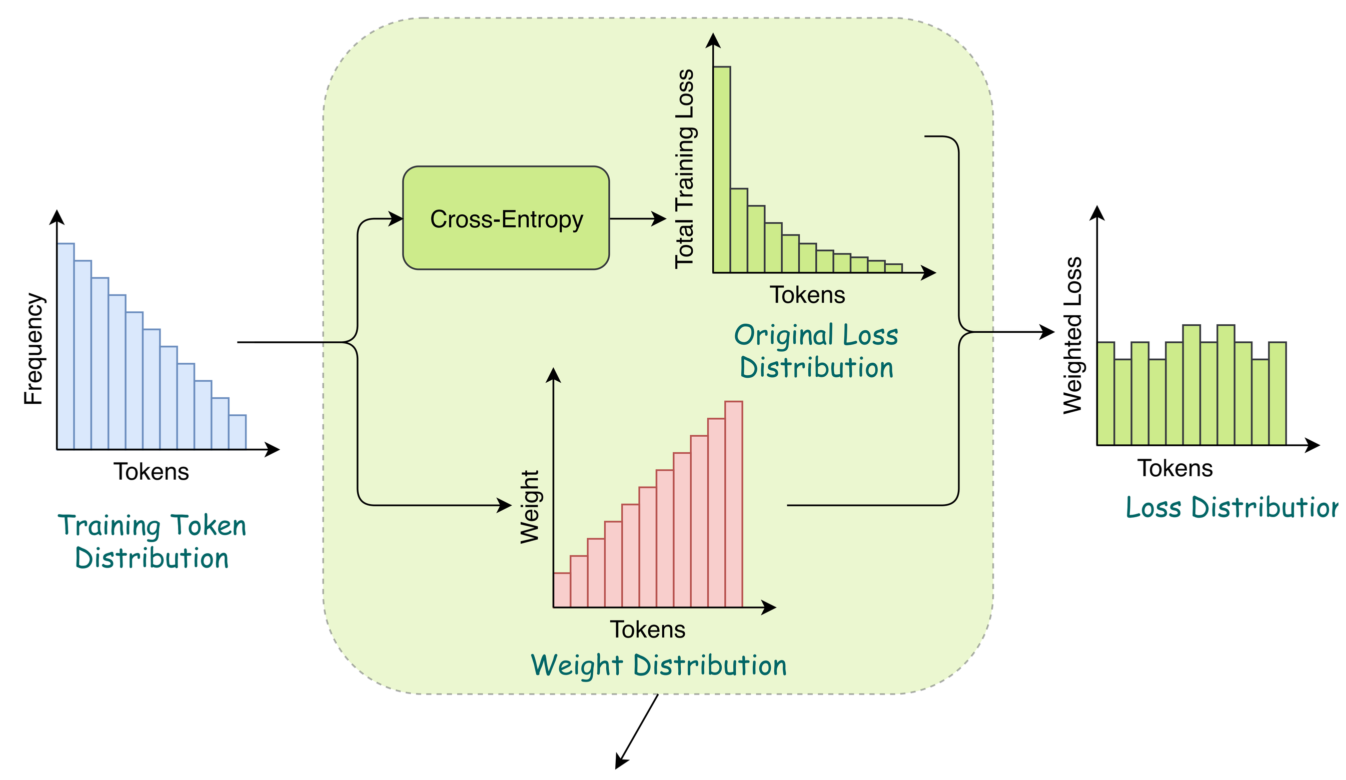
Why does it influence the response diversity?

Frequent Tokens + Language Model = Frequent Responses

Frequency-Aware Cross-Entropy Loss

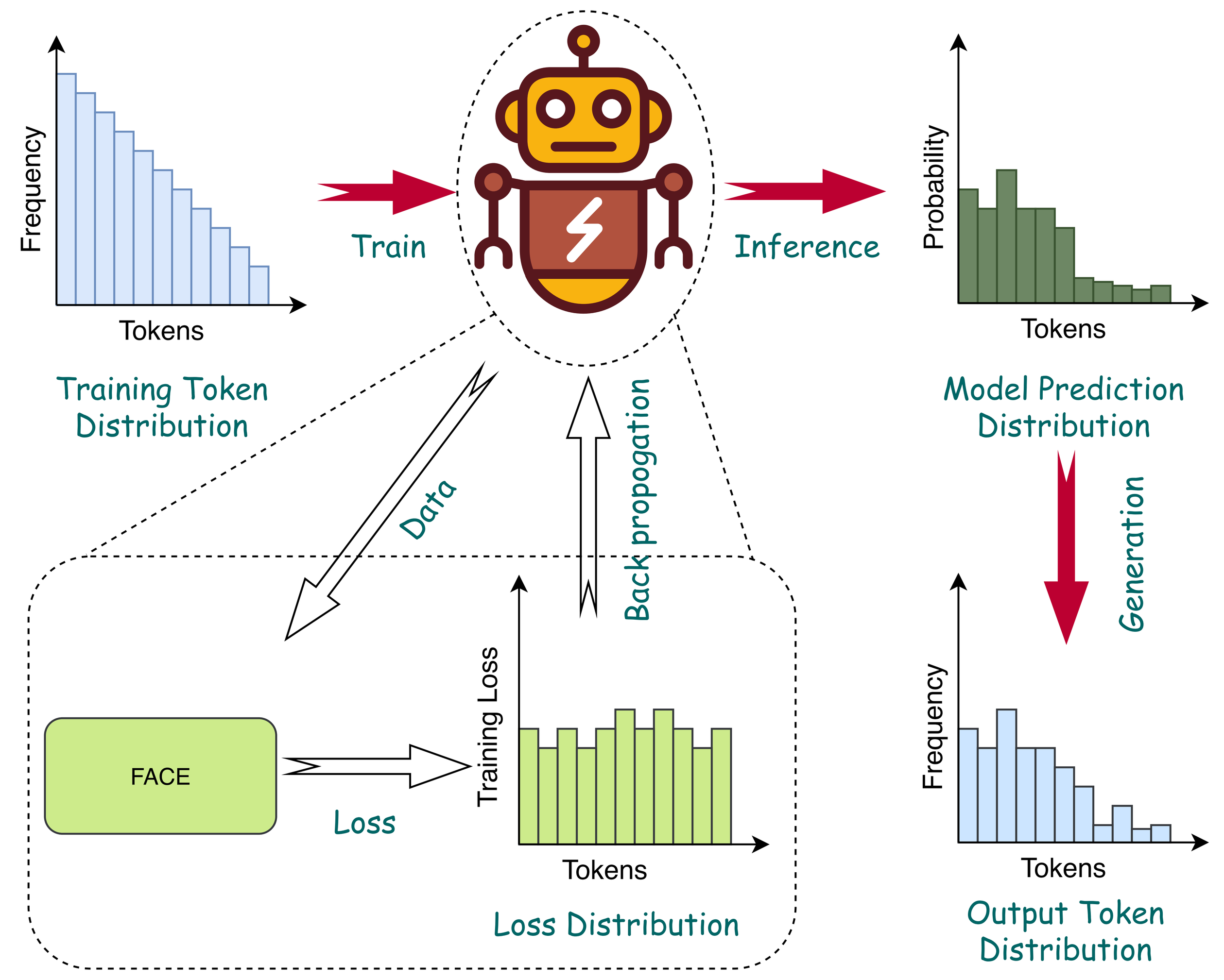
How do we make token generation more balanced? We balance model training loss: that's how model is directly influenced.

FACE loss function:



$$FACE(y_t) = w_i CE(y_t = c_i)$$

Model trained with FACE:



Experimental Results

Automatic & human evaluations on the Twitter dataset:

Table: Automatic evaluation results; d-n indicates distinct n-grams.

Model	d-1	d-2	BLEU
Seq2Seq	5.87	17.73	8.77
MMI-antiLM	**6.23**	18.21	6.51
MMI-bidi	5.42	15.16	8.20
MHAM	5.52	17.04	**8.96**
CMHAM	4.99	14.91	8.71
FACE	**6.23**	**24.18**	8.33

Table: Human evaluation results ($p\text{-}value < 0.05$ (*), < 0.01 (**), < 0.005 (***)).

Comparison	Win	Lose	Gain
FACE vs Seq2Seq	**38.61*****	21.54	17.07
FACE vs MMI-antiLM	**51.30*****	19.35	31.95
FACE vs MMI-bidi	**61.91*****	20.92	40.99
FACE vs MHAM	**50.93****	42.56	8.37
FACE vs CMHAM	**43.75***	38.85	4.90

Some examples:

History	Responses
Context: i need 4 mo followers .. **Message**: u ever watched criminal minds suspect behavior??	***FACE***: no i do n't think so .. i 'm watching it now .. ***Seq2Seq***: i do n't know what to say ..
Context: dont you dare steal my idea. i call shenanigans **Message**: Totally not stealing. Everyone collaborated on that idea.	***FACE***: well then you should have said it like that . ***MMI-antiLM***: come on . come on go back to you*
Context: I never watched Spongebob's movie tho, was it good? You know Hey Arnold? **Message**: yes it is good u should watch :D haha hey arnould with footballegg face :P	***FACE***: haha , i 'm watching it now , it 's so good : d ***MMI-bidi***: oh , that
Context: perhaps a recipe is in order? I have room in my fridge... ;) **Message**: trying this one for 1st time LOL http://bit.ly/nrFMIU start from bottom recipe & work up	***FACE***: sounds good ! i 'll be sure to check it out ! thanks for sharing the mocha cakes with the butter cakes ;) ***MHAM***: i 'm sure you 'll find a recipe for me !
Context: Don't laugh at my picture! That's not nice. **Message**: i'm laughing because you actually posted the pic. I believed you:-) lol	***FACE***: i 'm not a fan of the picture but it 's not that bad , it 's just a pic of me saying " i 'm a girl " haha : -) ***CMHAM***: i 'm not laughing at you .

Conclusion

- FACE loss function can effectively improve diversity and quality of responses
- FACE achieves improvements with minimum modifications to original Seq2Seq model, which makes it flexible to extend
- A limitation of FACE is that learning procedure is not as stable as cross-entropy, which increases difficulty of training

Source Code & Twitter

UNIVERSITY OF AMSTERDAM

TheWebConf 2019, San Francisco, The United States

{s.jiang, p.ren, c.monz, derijke}@uva.nl