Dialog Context Language Modeling with Recurrent Neural Networks

Bing Liu, Ian Lane
Carnegie Mellon University, liubing@cmu.edu, lane@cmu.edu

Abstract

We propose contextual language models that incorporate dialog level discourse information into language modeling. Previous works on contextual language model treat preceding utterances as a sequence of inputs, without considering dialog interactions. We design recurrent neural network (RNN) based contextual language models that specially track the interactions between speakers in a dialog. Experiment results on Switchboard Dialog Act Corpus show that the proposed model outperforms conventional single turn based RNN language model by 3.3% on perplexity. The proposed models also demonstrate advantageous performance over other competitive contextual language models.

Introduction / Motivation

- RNN language model becomes an increasingly popular choice in many areas such as speech recognition and spoken dialog systems.
- Contextual information, such as topic of the dialog and the dialog act, can be encoded and added to RNN to make the language model context-aware.
- Language models using document level context have been proposed [1, 2] by encoding text before the target sentence as context vectors.
- Such document context models may not be suitable for modeling turn-taking in a dialog, as they simply treat preceding sentences as a sequence of inputs, without modeling interactions between speakers.
- We address this problem by proposing contextual RNN language models that specially track the interactions between speakers in a dialog.

Proposed Methods

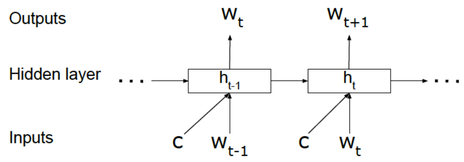
Dialog Context Language Model

Let $D = (U_1, U_2, \dots, U_K)$ be a dialog of K turns involving two speakers. The kth turn $U_k = (w_1, w_2, \dots, w_{T_k})$ has T_k words.

Probability of the target turn U_k given dialog context:

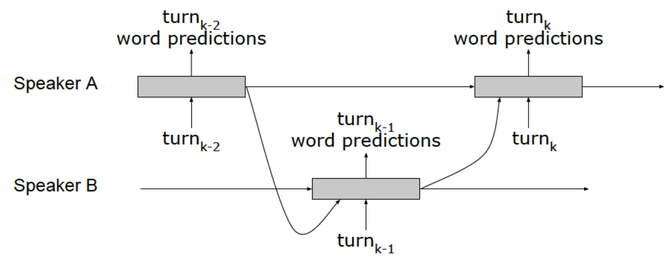
$$P(\mathbf{U}_k|\mathbf{U}_{<k}) = \prod_{t=1}^{T_k} P(w_t^{\mathbf{U}_k}|w_{<t}^{\mathbf{U}_k}, \mathbf{U}_{<k})$$

Contextual RNNLM



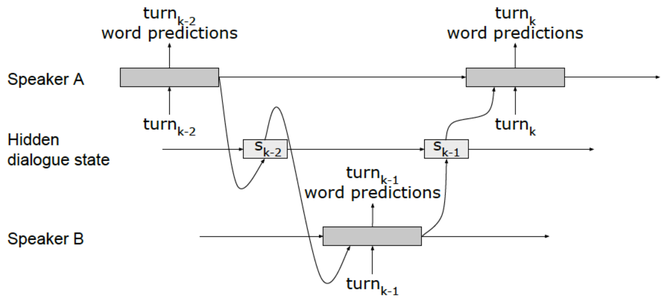
$$h_t = \text{RNN}(h_{t-1}, [w_t, c])$$

Interactive Dialog Context LM (IDCLM)



$$c^{\mathbf{U}_k} = h_{T_{k-1}}^{\mathbf{U}_{k-1}},\ h_0^{\mathbf{U}_k} = h_{T_{k-2}}^{\mathbf{U}_{k-2}}$$

External State Interactive Dialog Context LM (ESIDCLM)



$$s_{k-1} = \text{RNN}_{ES}(s_{k-2}, h_{T_{k-1}}^{\mathbf{U}_{k-1}})$$

$$c^{\mathbf{U}_k} = s_{k-1},\ h_0^{\mathbf{U}_k} = h_{T_{k-2}}^{\mathbf{U}_{k-2}}$$

Experiments and Results

Data set: Switchboard Dialog Act Corpus (SwDA)

- Train (sw00 to sw09): 98.7K turns (190.0K utterances)
- Dev (sw11 to sw13): 5.7K turns (11.3K utterances)
- Test (sw10): 11.9K turns (22.2K utterances)
- Vocab size: 10K

Table 1: Perplexities on SwDA corpus with different dialog context turn sizes (K)

Model	K=1	K=2	K=3	K=5
5-gram KN	65.7	-	-	-
Single-Turn-RNNLM	60.4	-	-	-
BoW-Context-RNNLM	-	59.6	59.2	58.9
DRNNLM	-	60.1	58.6	59.1
CCDCLM	-	63.9	61.4	62.2
IDCLM	-	-	58.8	58.6
ESIDCLM	-	-	**58.4**	**58.5**
DACLM	-	58.2	57.9	58.0

RNNLM with TRUE dialog act as context

Table 2: Perplexity relative change (%) per POS tag

POS Tag	IDCLM	ESIDCLM	DACLM
PRP	-16.8	-5.8	-10.1
IN	-2.0	-5.5	-1.8
RB	-4.1	-8.9	-4.3
NN	13.4	8.1	2.3
UH	-0.4	7.7	-9.7

Consistent performance Gain for pronouns, prepositions, and adverbs

Table 3: Perplexity relative change (%) per Dialog Act tag

DA Tag	IDCLM	ESIDCLM	DACLM
Statement-non-opinion	-1.8	-0.5	-1.6
Acknowledge	-2.6	11.4	-16.3
Statement-opinion	4.9	-0.9	-1.0
Agree/Accept	14.7	2.7	-15.1
Appreciation	0.7	-3.8	-6.5

Conclusions

- We propose two dialog context language models that with special design to model dialog interactions.
- Our proposed models outperform conventional RNN language model by 3.3% on SwDA corpus.
- The proposed models show advantageous performance over several competitive contextual language models.

References

[1] Tian Wang and Kyunghyun Cho, "Larger-Context Language Modelling with Recurrent Neural Network," in *ACL*, 2016
[2] Yangfeng Ji et al., "Document context language models," in *ICLR Workshop*, 2016