ServiceNet: End to End Dialogue System for Technical Support

Madhulima Pandey, and David Uvalle
CS230 Deep Learning, Stanford University

Abstract

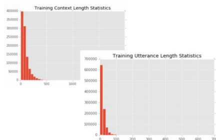
IT support is a manually intensive task where support staff responds to a variety of user issues and request for information. Current production grade dialogue systems require of multiple components to work and are focused on slot filling of predesigned conversation flows. We propose an end-to-end dialogue using deep learning that could produce highly accurate responses and would be easier and cheaper to develop and maintain.

Data and Features

We used Ubuntu Dialog Corpus created from a collection of logs from Ubuntu-related chat rooms for technical support on the IRC network.

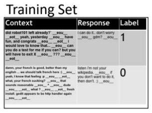
Dataset:

- # dialogues: 930,000
- # utterances: 7,100,000
- # words total: 100,000,000
- Min/avg # turns: 3/7.71
- Avg. # words: 10.34
- Vocab size: 91,000

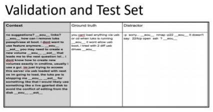


Data Preprocessing: We process raw data into training data that contains context, response utterance and a flag label that indicates if the response was the actual next utterance for the given context. Validation and test set contain context, ground truth and multiple distractors

Training Set

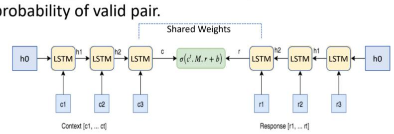


Validation and Test Set

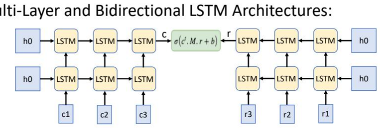


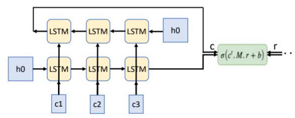
Models

- Dual Encoder Architecture with Siamese Recurrent Networks.
- Context and Responses inputs to each branch of the network.
- Embeddings of the Context and Responses used to calculate probability of valid pair.

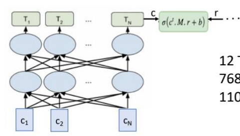


- Probability of Valid Pair: $p(\text{flag}=1|c,r,M)=\sigma(c^T Mr+b)$
- Loss Function: $\mathcal{L}=-\sum_n \log p(\text{flag}_n|c_n,r_n,M)+\frac{\lambda}{2}||\theta||_F^2$
- Multi-Layer and Bidirectional LSTM Architectures:





- BERT - BERT: Pre-trained Deep Bidirectional Transformers



12 Transformer Blocks
768 Hidden states
110M variables

Results

Network Type	State Size	Num Layers	Bidirectional	Pre-trained	Recall @1	Recall @2	Recall @5
LSTM	256	1	No	No	0.526	0.708	0.923
LSTM	256	1	No	Yes	0.519	0.708	0.919
LSTM	768	1	No	Yes	0.499	0.683	0.914
LSTM-Multi	128	2	No	Yes	0.494	0.689	0.911
LSTM-Multi	128	3	No	Yes	0.483	0.675	0.911
LSTM-Bidi	128	1	Yes	Yes	0.512	0.696	0.915
LSTM-Bidi	128	1	Yes	No	0.537	0.712	0.920
BERT	768	12	Yes	Yes	0.551	0.738	0.927

Discussion and Future Work

Observations

- LSTMs substantially help improve performance over baseline TF-IDF results (R@1 41%, R@5 71%)
- Bidirectional LSTM helps improve accuracy more than increasing state size or increasing number of layers
- Pretrained embeddings did not help in increasing accuracy
- Increasing state size beyond a limit can results in decrease of accuracy. State size 256 yields best results

Future Work

- Fine-tune pre-trained network (BERT, ELMO) with UDC corpus to improve accuracy
- Use knowledge graph and additional context information

Acknowledgements – Mentor – Pedro Garzon, CS230 staff

References

- Bordes, et al. Learning end-to-end goal-oriented dialog
- Lowe, et al. Training end-to-end ubuntu dialogue corpus
- Devin, et al. BERT: Pre-training of Deep Bidi Transformers