Aspect Based Sentiment Analysis using Deep Memory Networks

Andrei Vacariu
Wasifa Chowdhury
Stanley Gan

Outline of Presentation

- Review of task
- Network architecture
- Approaches
- Results and comparison
- Future work
- References

Review of Our Task

Semeval Task 4, Subtask 4: Aspect Category Polarity

- Given a set of categories and a review, determine the polarity of each category
- Example
  - Given categories: Food, price
  - Given review: “The restaurant was too expensive, but the menu was great”
  - Food: Positive, Price: Negative
- 4 classes of sentiment: Positive, Neutral, Negative, Conflict
- Dataset
  - Restaurant reviews: 3041 training data, 100 test data
  - Laptop reviews: 3045 training data, 100 test data
- Evaluation
  - Accuracy: number of correctly predicted polarity labels/ total number of labels

Network Architecture: Memory Networks

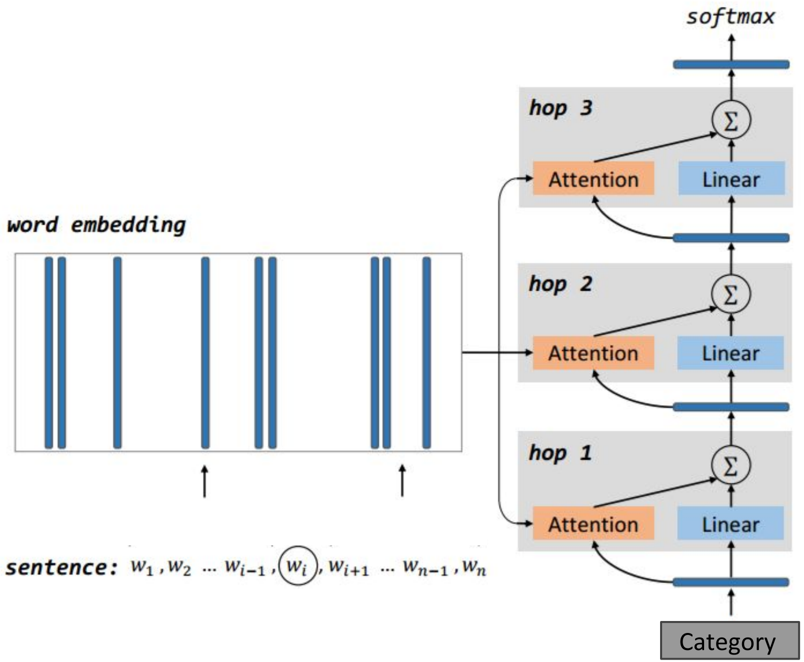


Figure 1: An illustration of our deep memory network with three computational layers (hops) for aspect level sentiment classification.

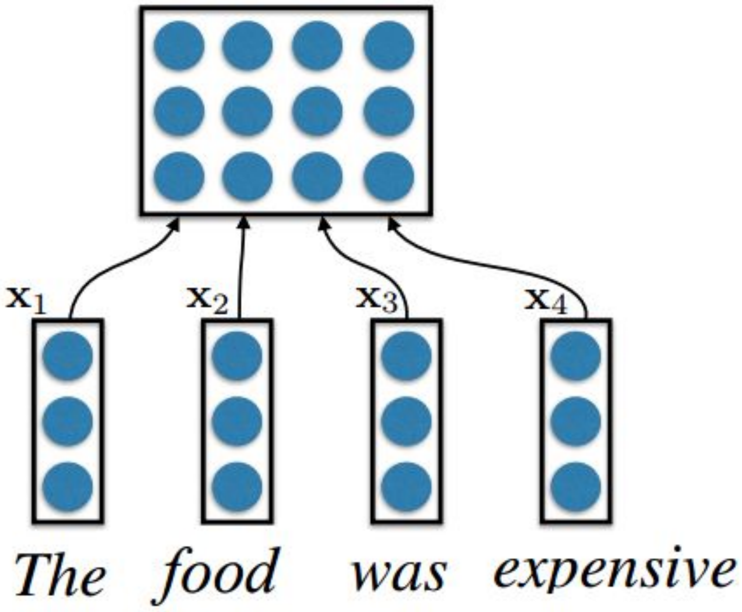
- Aspect category instead of aspect term
- Features of Memory Networks
  - Explicit memory
  - Attention layers and hops
- Using pre-trained GloVe vectors for initial word embeddings

Our Approaches

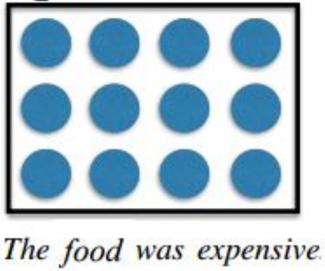
- Bidirectional LSTM for input representations
- Incorporate aspect into memory vectors
- Linear mapping from pre-trained GloVe vectors to context vectors
- Local attention mechanism
- Incorporate aspect into hop output

Input Representations

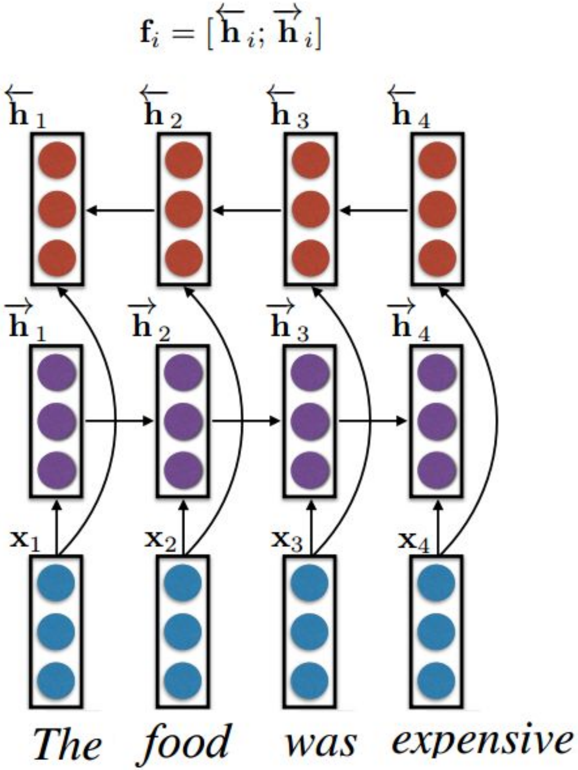
$$\mathbf{f}_i = \mathbf{x}_i$$

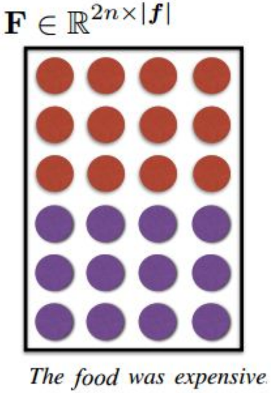


$\mathbf{F} \in \mathbb{R}^{n \times |\boldsymbol{f}|}$

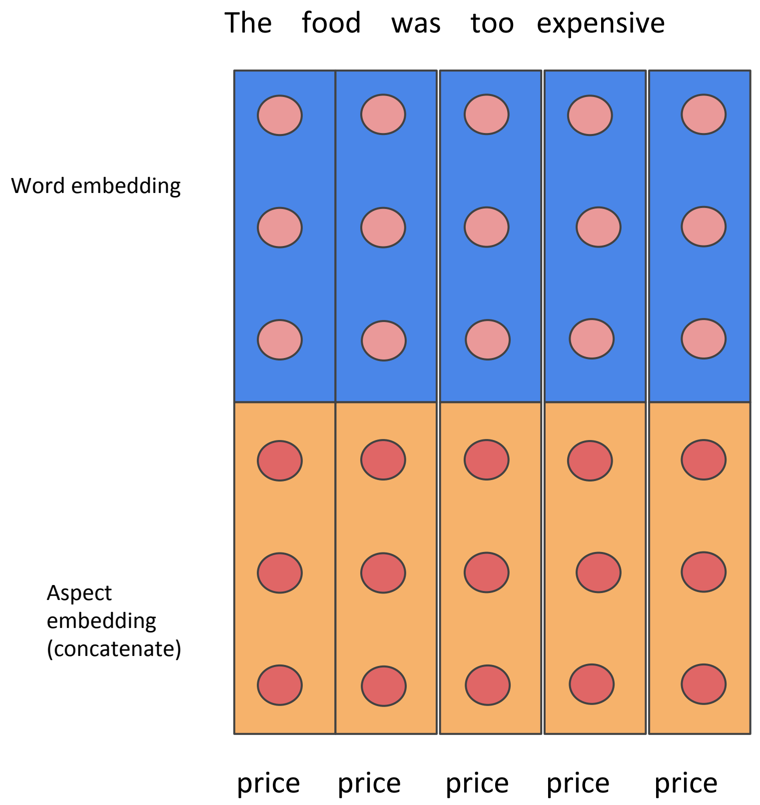


Input Representations: Bidirectional LSTM



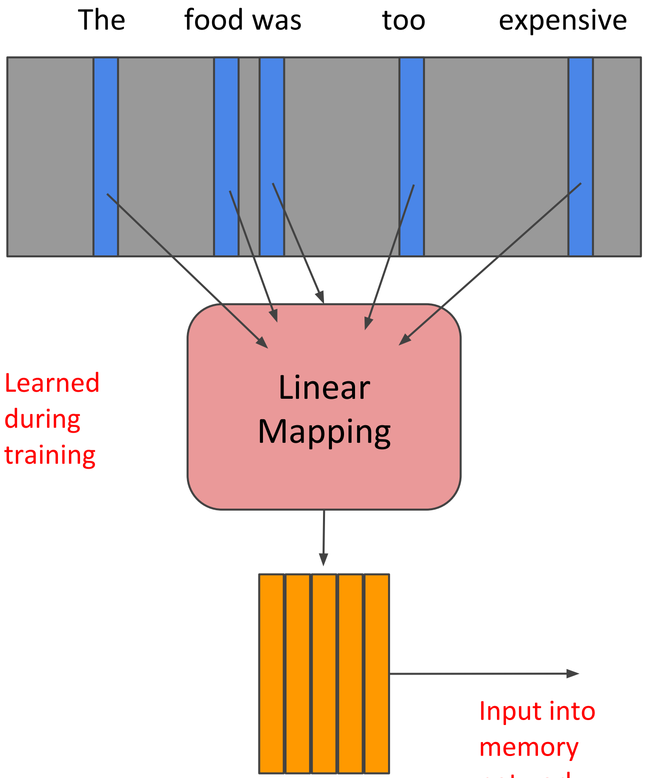


Input Representations: Incorporating Aspect



Linear Mapping from GloVe to Context

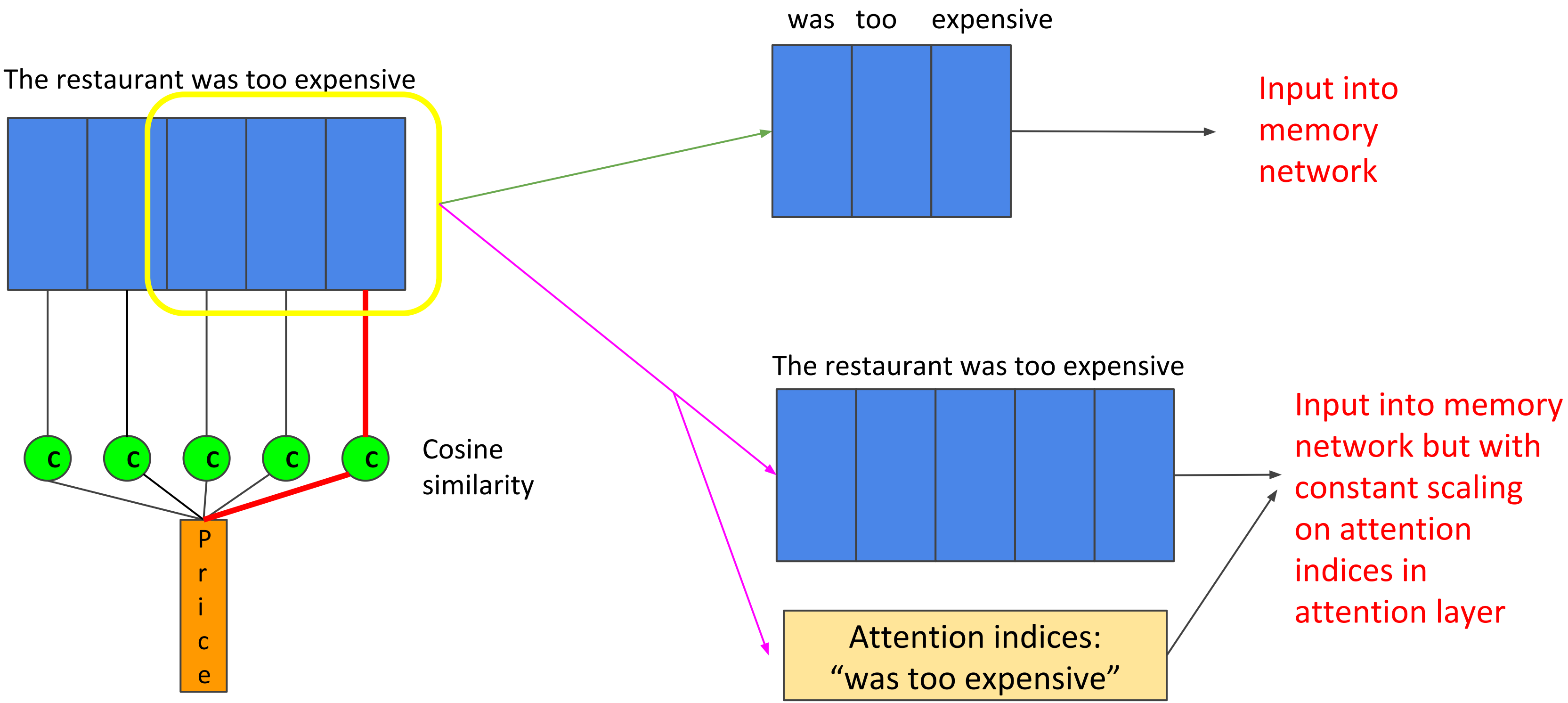
Pre-trained GloVe vectors



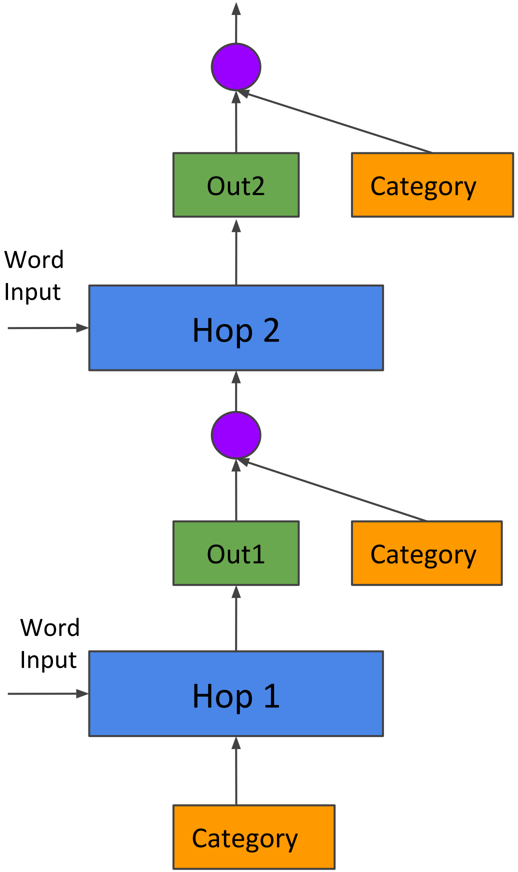
Learned during training

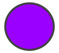
Input into memory network

Local Attention Mechanism



Incorporating Aspect to Hop Output




= concatenate/linear combination

Concatenate

- Concatenate hop output and category
- Linear layer

Linear Combination

- $New\_output_k = \alpha_k * Category + (1 - \alpha_k) * Out_k$
- α_k is learned

Results on Different Configurations

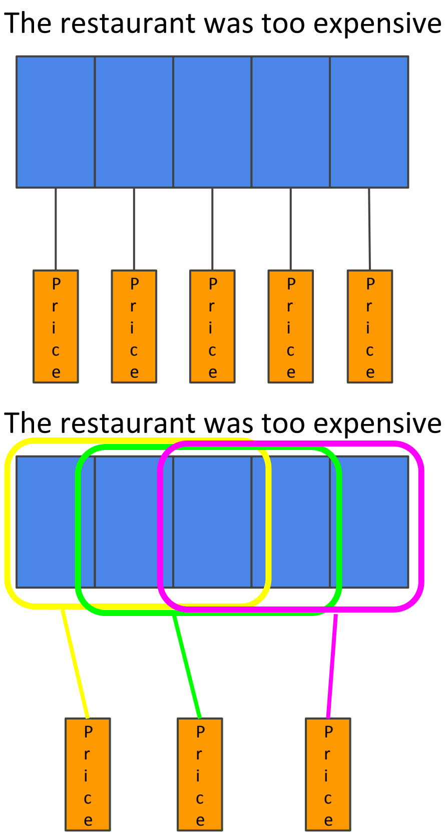
	Laptop			Restaurant		
	2 classes	**3 classes**	**4 classes**	**2 classes**	**3 classes**	**4 classes**
BiLSTM	85.93	72.41	66.53	88.93	83.7	75.59
BiLSTM+AC	85.93	71.68	68.28	88.80	83.7	78.42
BiLSTM + AS=2	**87.24**	72.85	68.44	88.41	84.04	79.2
BiLSTM+AC+AS=2	**87.24**	**73.44**	67.81	88.54	**84.8**	**79.49**
LA	86.46	69.53	64.69	88.15	79.46	75.59
LA+AC	84.63	69.53	64.38	86.59	80.35	75.10
HopLC + LA	73.44	23.04	63.75	75.78	69.08	64.06
HopLC + LA + AC	73.43	-	61.56	75.78	69.08	64.06
HopConcat + LA	85.93	69.14	**68.81**	88.93	81.03	71.93
HopConcat + LA + AC	85.16	69.72	63.75	88.41	81.14	72.66
BiLSTM + HopLC + AS=2	-	20.31	51.72	78.71	69.04	64.06
BiLSTM + HopConcat + AS=2	86.72	70.70	63.96	**89.45**	81.47	75.10

Our Best Configuration vs. State of the Art

2 classes		
	Laptop	**Restaurant**
Our best	87.24	89.45
ATAE-LSTM	**87.6**	**90.9**
AE-LSTM	87.4	89.6

3 classes		
	Laptop	**Restaurant**
Our best	**73.44**	**84.8**
ATAE-LSTM	68.7	77.2
AE-LSTM	68.9	76.6

Future Potential Approaches



Semantic relatedness between category and each word

Semantic relatedness between category and multiple words

1. Use GloVe vectors trained on twitter data
2. ConvNet for attention layer

References

- Tang, Duyu, Bing Qin, and Ting Liu. "Aspect level sentiment classification with deep memory network." *arXiv preprint arXiv:1605.08900* (2016).
- Sukhbaatar, Sainbayar, Jason Weston, and Rob Fergus. "End-to-end memory networks." In *Advances in neural information processing systems*, pp. 2440-2448. 2015.
- Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." *arXiv preprint arXiv:1409.0473* (2014).
- Tang, Duyu, Bing Qin, Xiaocheng Feng, and Ting Liu. "Effective LSTMs for Target-Dependent Sentiment Classification." *arXiv preprint arXiv:1512.01100* (2015).
- Wang, Yequan, Minlie Huang, Li Zhao, and Xiaoyan Zhu. "Attention-based LSTM for Aspect-level Sentiment Classification."
- Weston, Jason, Sumit Chopra, and Antoine Bordes. "Memory networks." *arXiv preprint arXiv:1410.3916* (2014).
- Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning. "Effective approaches to attention-based neural machine translation." *arXiv preprint arXiv:1508.04025* (2015).
- Socher, Richard, Alex Perelygin, Jean Y. Wu, Jason Chuang, Christopher D. Manning, Andrew Y. Ng, and Christopher Potts. "Recursive deep models for semantic compositionality over a sentiment treebank." In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, vol. 1631, p. 1642. 2013.
- Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. "Glove: Global Vectors for Word Representation." In *EMNLP*, vol. 14, pp. 1532-1543. 2014.
- Dataset: SemEval(Semantic Evaluation) 2014
- Partially implemented TensorFlow code for Deep Memory Networks: https://github.com/ganeshjawahar/mem_absa

Thank you for listening

Q&A?