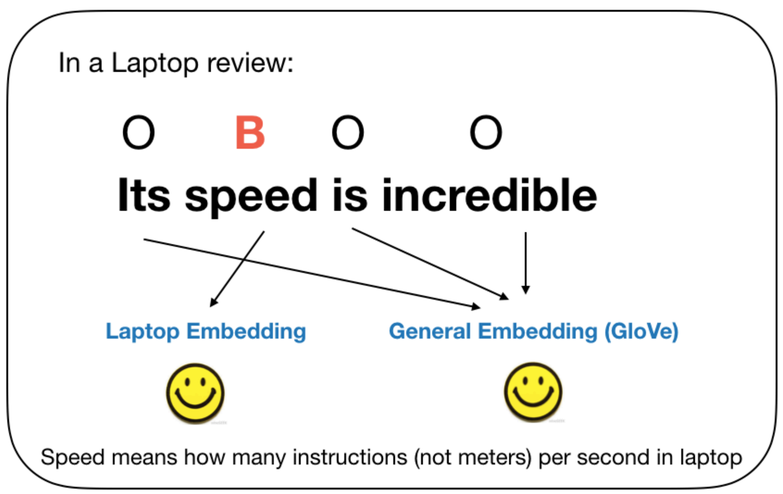
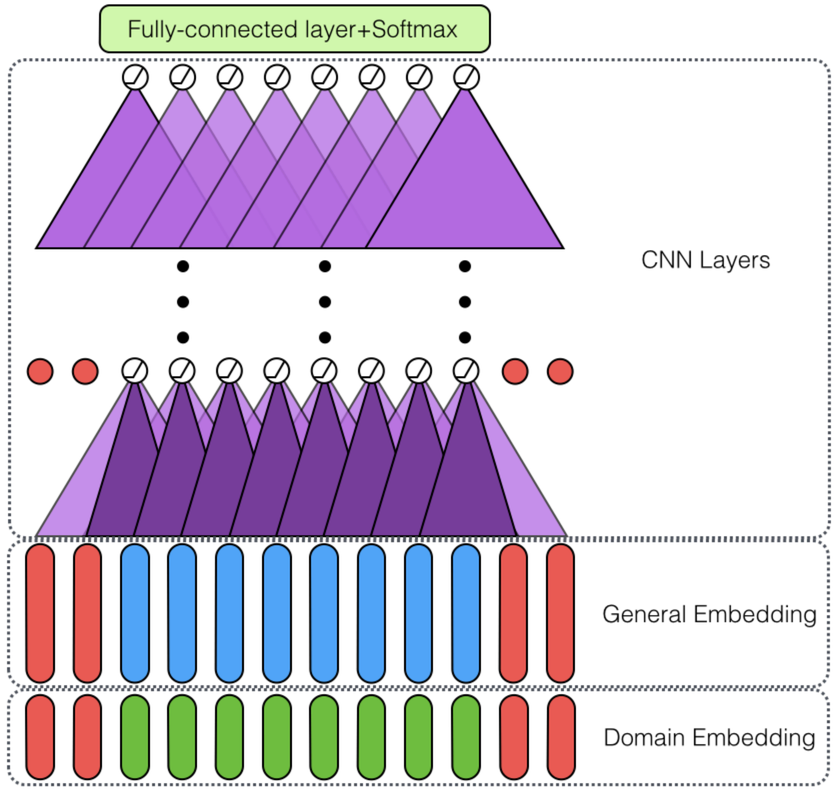
Double Embeddings and CNN-based Sequence Labeling for Aspect Extraction

Hu Xu[1], Bing Liu[1], Lei Shu[1], Philip S. Yu[1,2] {hxu48, liub, lshu3, psyu}@uic.edu

[1]Department of Computer Science, University of Illinois at Chicago
[2]Institute for Data Science, Tsinghua University

Aspect Extraction is an important task in fine-grained sentiment analysis. We propose a simple and fast approach without using any sophisticated features and models.
The contributions are in 2 folds:
Double embedding: we use two types of pre-trained embeddings for aspect extraction: general purpose embeddings and domain specific embeddings.
CNN: we use CNN for sequence labeling, which is parallel and faster than serial LSTM. We adapt CNN (e.g., drop max-pooling layer) to get better results.





Architecture of DE-CNN: red vectors are padding vectors

The r-th CNN filter for the i-th word in layer l

$$x_{i,r}^{(l+1)} = \max\left(0, \left(\sum_{j=-c}^{c} w_{j,r}^{(l)} x_{i+j}^{(l)}\right) + b_r^{(l)}\right), \quad (1)$$

Description	Training #S./#A.	Testing #S./#A.
SemEval-14 Laptop	3045/2358	800/654
SemEval-16 Restaurant	2000/1743	676/622

Table 1: Dataset description with the number of sentences(#S.) and number of aspect terms(#A.)

Model	Laptop	Restaurant
CRF	74.01	69.56
IHS_RD	74.55	-
NLANGP	-	72.34
WDEmb	75.16	-
LSTM	75.25	71.26
BiLSTM-CNN-CRF	77.8	72.5
RNCRF	78.42	-
CMLA	77.80	-
MIN	77.58	73.44
GloVe-CNN	77.67	72.08
Domain-CNN	78.12	71.75
MaxPool-DE-CNN	77.45	71.12
DE-LSTM	78.73	72.94
DE-OOD-CNN	80.21	74.2
DE-Google-CNN	78.8	72.1
DE-CNN-CRF	80.8	74.1
DE-CNN	**81.59***	**74.37***

Table 2: Comparison results in F_1 score: numbers in the third group are averaged scores of 5 runs. * indicates the result is statistical significant at the level of 0.05.

The implementation of the paper is available at https://www.cs.uic.edu/~hxu/