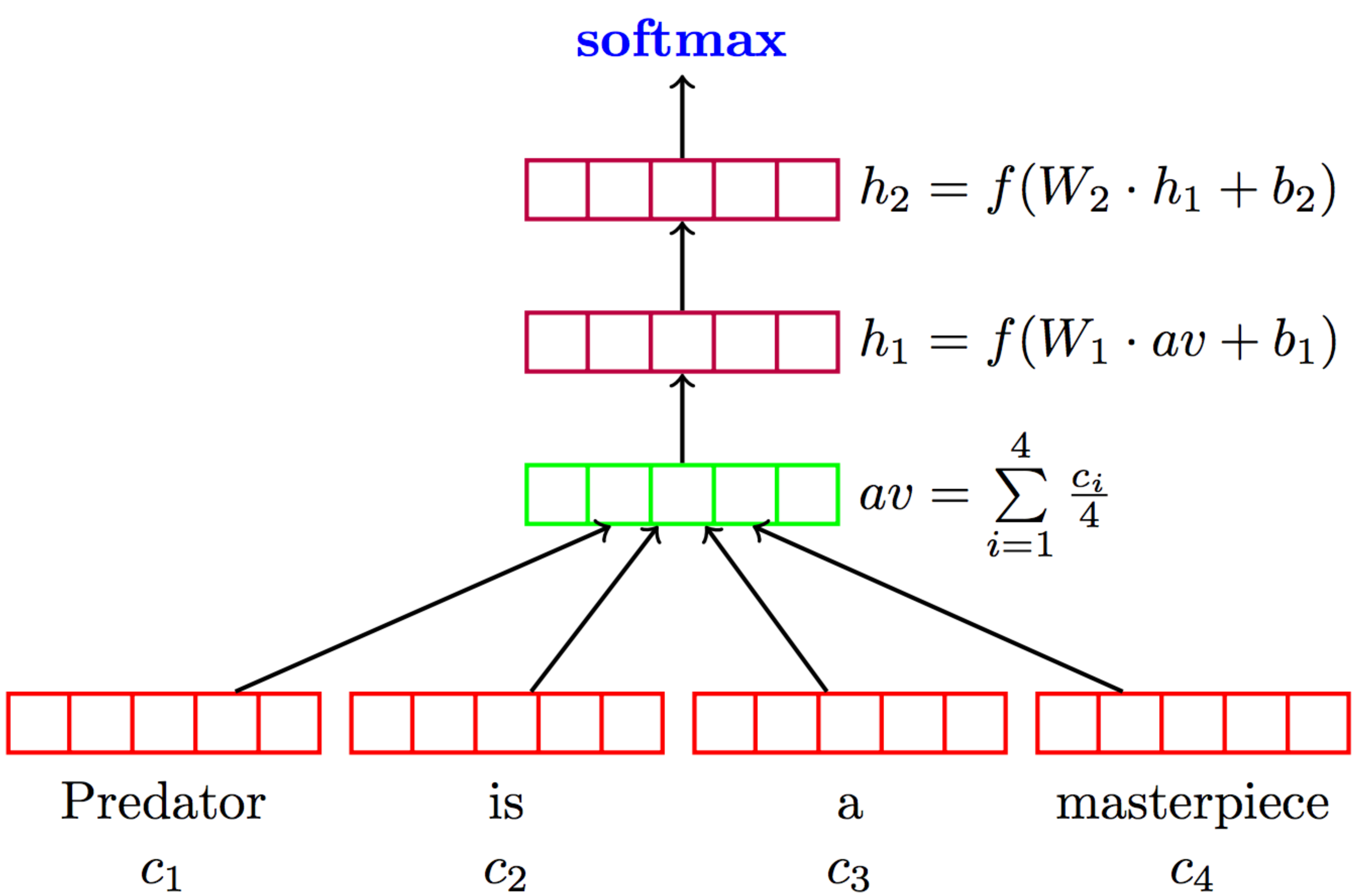
Applying Embeddings

- First layer of your network: map from word indices to embeddings
- Approach 1: learn these embeddings as parameters from your data
  - Often works pretty well
- Approach 2: initialize word embeddings using GloVe, keep fixed
  - Faster because no need to update these parameters
- Approach 3: initialize word embeddings GloVe, fine-tune
  - Works best for some tasks
- Can also evaluate embeddings intrinsically on tasks like word similarity

Deep Averaging Networks

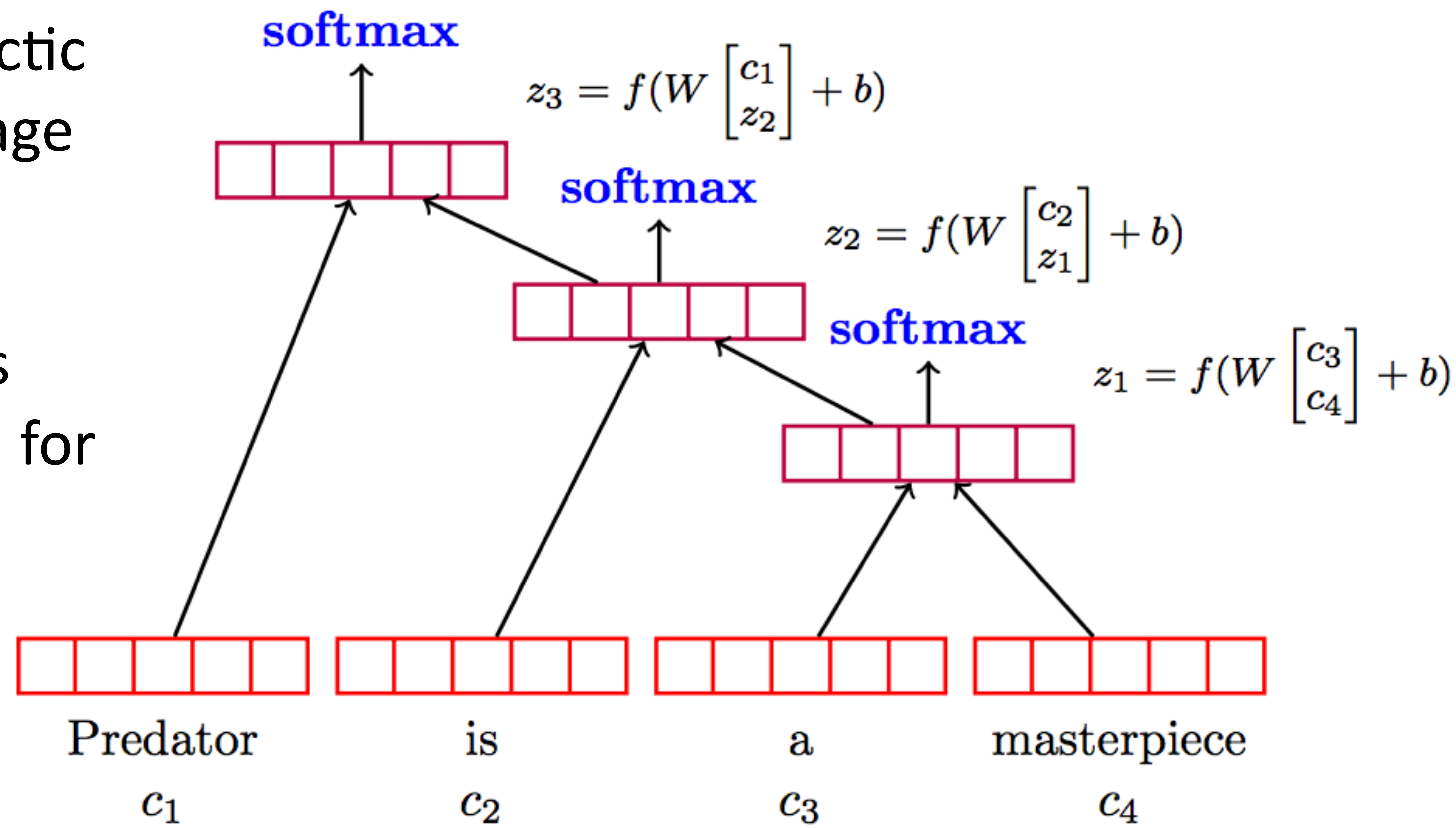
- Deep Averaging Networks: feedforward neural network on average of word embeddings from input



Iyyer et al. (2015)

Deep Averaging Networks

- Contradicts a widely-held view that we need to model syntactic structure to represent language
- Simple averaging can work as well as syntactic composition for some problems!



Iyyer et al. (2015)

Deep Averaging Networks

Model	RT	SST fine	SST bin	IMDB	Time (s)
DAN-ROOT	—	46.9	85.7	—	**31**
DAN-RAND	77.3	45.4	83.2	88.8	136
DAN	80.3	47.7	86.3	89.4	136
NBOW-RAND	76.2	42.3	81.4	88.9	91
NBOW	79.0	43.6	83.6	89.0	91
BiNB	—	41.9	83.1	—	—
NBSVM-bi	79.4	—	—	91.2	—
RecNN*	77.7	43.2	82.4	—	—
RecNTN*	—	45.7	85.4	—	—
DRecNN	—	49.8	86.6	—	431
TreeLSTM	—	**50.6**	86.9	—	—
DCNN*	—	48.5	86.9	89.4	—
PVEC*	—	48.7	87.8	**92.6**	—
CNN-MC	**81.1**	47.4	**88.1**	—	2,452
WRRBM*	—	—	—	89.2	—

No pretrained embeddings

Bag-of-words

Tree-structured neural networks

Iyyer et al. (2015)

Wang and Manning (2012)

Kim (2014)

Iyyer et al. (2015)

Deep Averaging Networks

Sentence	DAN	DRecNN	Ground Truth
who knows what exactly godard is on about in this film, but his words and images do n't have to add up to mesmerize you.	positive	positive	positive
it's so good that its relentless, polished wit can withstand not only inept school productions, but even oliver parker's movie adaptation	negative	positive	positive
too bad, but thanks to some lovely comedic moments and several fine performances, it's not a total loss	negative	negative	positive
this movie was not good	negative	negative	negative
this movie was good	positive	positive	positive
this movie was bad	negative	negative	negative
the movie was not bad	negative	negative	positive

- Will return to compositionality with syntax and LSTMs

Iyyer et al. (2015)