Word Embedding

One-hot

Word2Vec

Train Method

- Skip-Gram

  Given the centre word, predict the context

- Continous Bag Of Word(CBOW)

  Given the context, predict the conter word

GloVe (Global Vectors for Word Representation)

Co-ocurrence Matrix

$$X_{i,j} = \frac{n_{i,j}}{d_{i,j}}$$

Objective Function

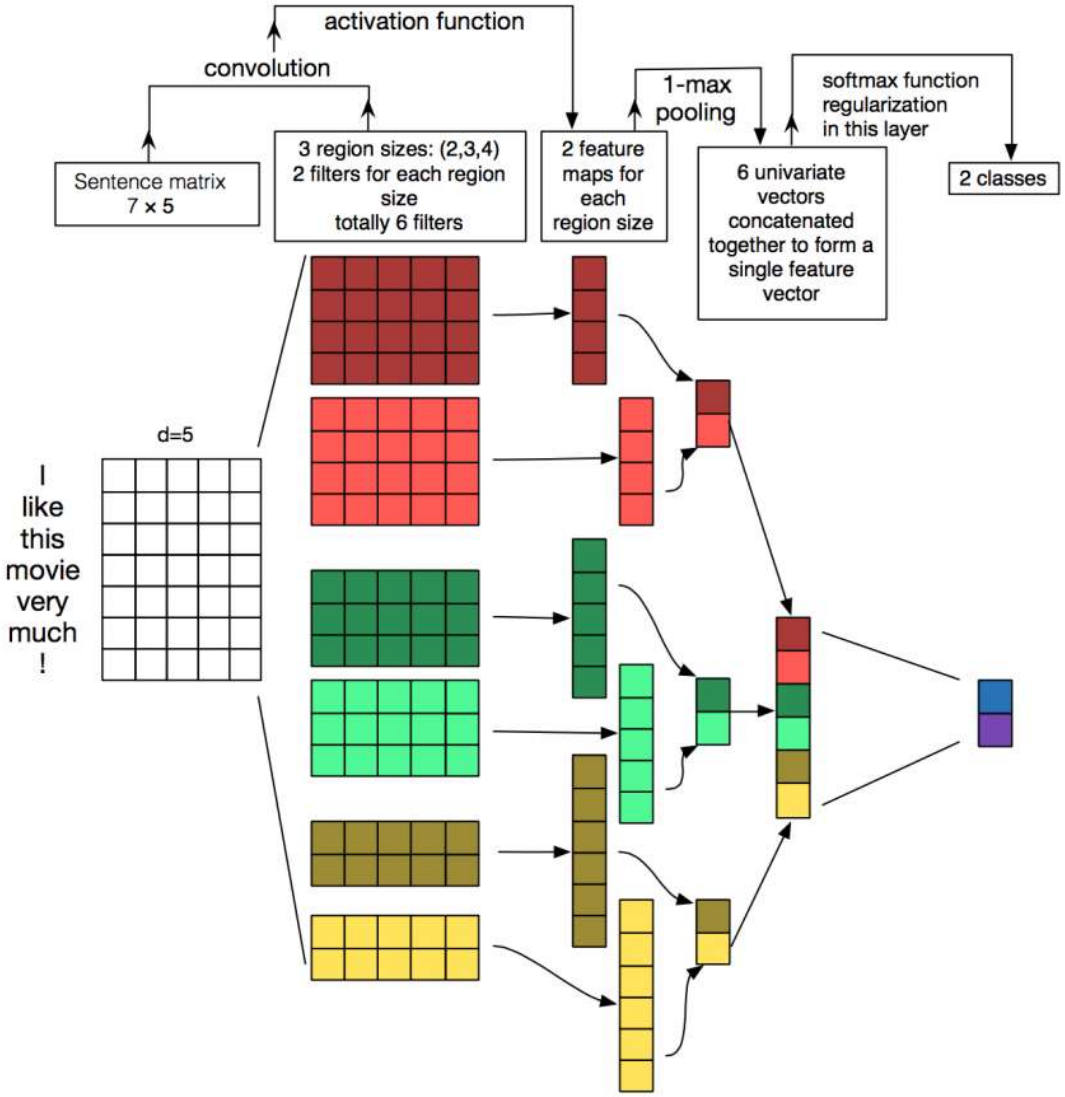
$$W_i^T W j + b_i + b_j = log(X_{i,j})$$

Loss Function

$$J = \sum_{i,j=1}^{V} f(X_{i,j}) * (W_i^T W_j + b_i + b_j - log(X_{i,j}))^2$$

Sentence Representation

Concatenate Word Embedding



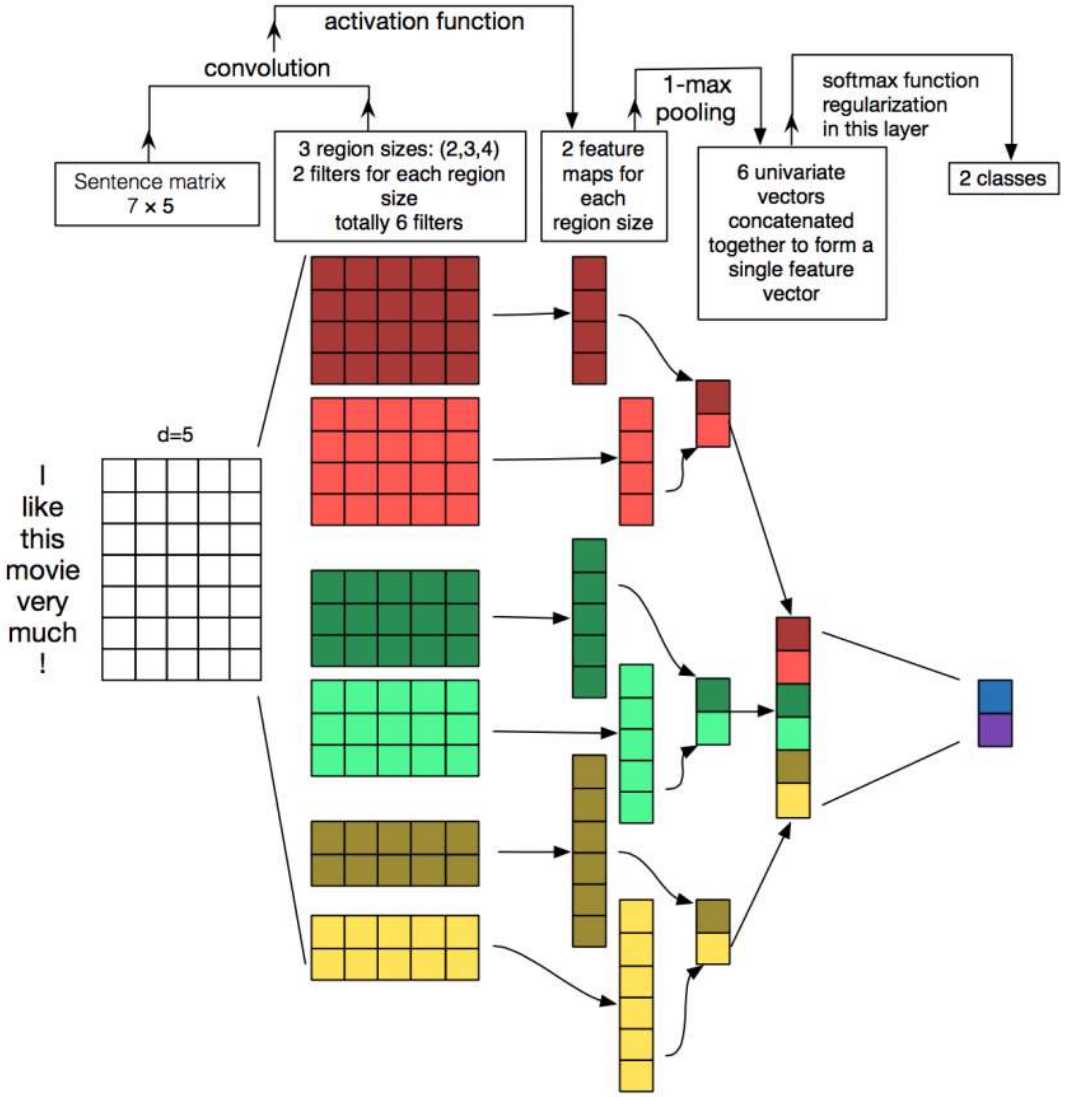
Character-Level

FastText

$$V_{sent} = \frac{\sum_{V_i \in W_{Sen}} V_i}{N}$$

NLP with DNN

Construction


activation function
convolution
1-max pooling
softmax function regularization in this layer
Sentence matrix 7 × 5
3 region sizes: (2,3,4) 2 filters for each region size totally 6 filters
2 feature maps for each region size
6 univariate vectors concatenated together to form a single feature vector
2 classes
d=5
I
like
this
movie
very
much
!

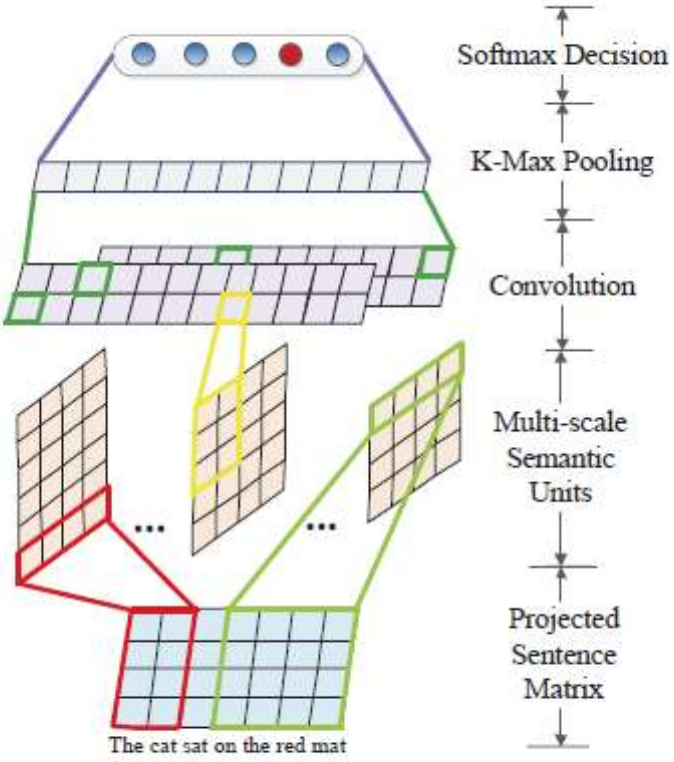
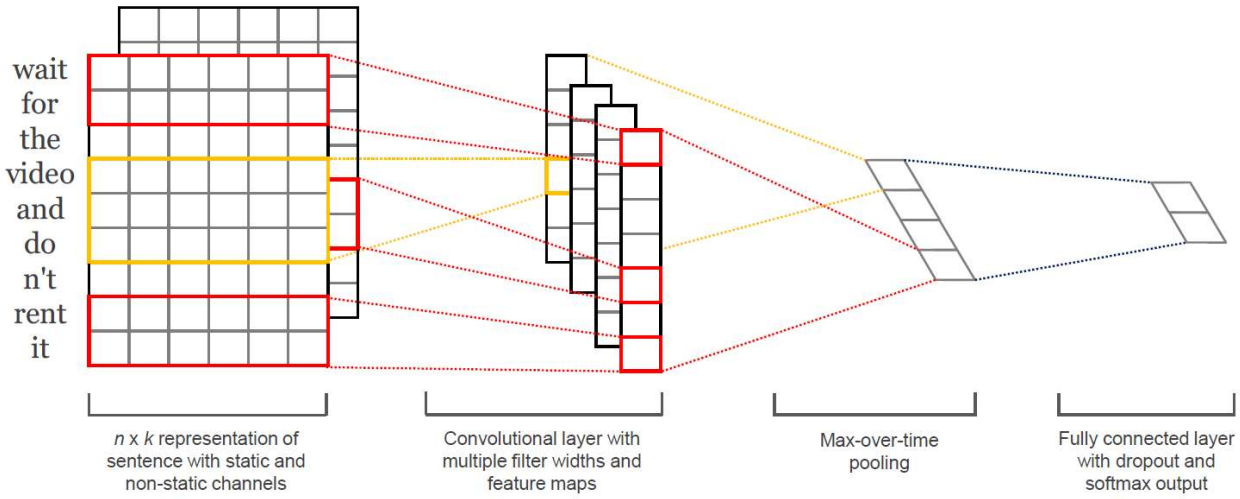


Figure 2: Architecture for short text modeling



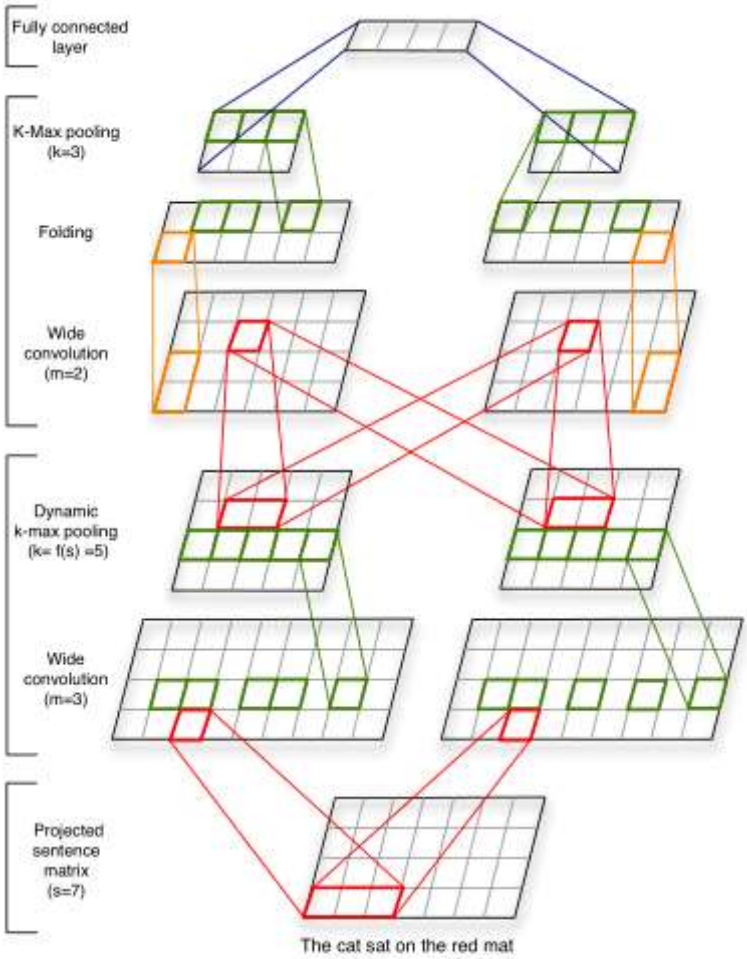


Figure 3: A DCNN for the seven word input sentence. Word embeddings have size $d = 4$. The network has two convolutional layers with two feature maps each. The widths of the filters at the two layers are respectively 3 and 2. The (dynamic) k-max pooling layers have values k of 5 and 3.

$$k_l = max(k_{top}, \lceil \frac{L-l}{L} s \rceil)$$

Multichannel

- more than one kind of word embedding
  - one-hot
  - Word2Vec
  - GloVe
- same sentence represented in different languages

Comparation of CNN and RNN for NLP

- CNNs and RNNs provide complementary information for text classification tasks. Which architecture performs better depends on how important it is to semantically understand the whole sequence.

- Learning rate changes performance relatively smoothly, while changes to hidden size and batch size result in large fluctuations.

References

- [1] Semantic Clustering and Convolution Neural Network for Short Text Categorization
- [2] Convolution Neural Network for Sentence Classification
- [3] Comparative Study of CNN and RNN for Natural Language Processing
- [4] http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/#more-348