IN5550: Neural Methods in Natural Language Processing
Sub-lecture 5.5
RNN practicalities

Andrey Kutuzov, Roxana Pop

University of Oslo



Contents



1. RNN tips and tricks

RNN tips and tricks



Variable Length of Input Sequence

- Although RNNs in principle are well-defined for inputs of variable length,
- in practice, padding to fixed length is required for efficiency (batching);
- actually not too much 'waste'; but can be beneficial to bin by length.

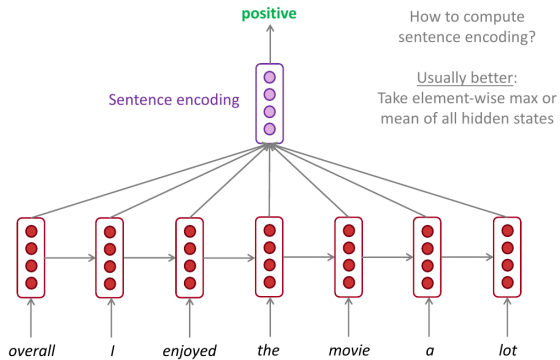
Evaluation

- Accuracy is a common metric for tagging; fixed number of predictions;
- for most inputs, very large proportion of padding tokens (and labels);
- trivial predictions will inflate accuracy scores; detrimental to learning?
- Can define custom accuracy functions; mask padding tokens predictions.

RNN tips and tricks

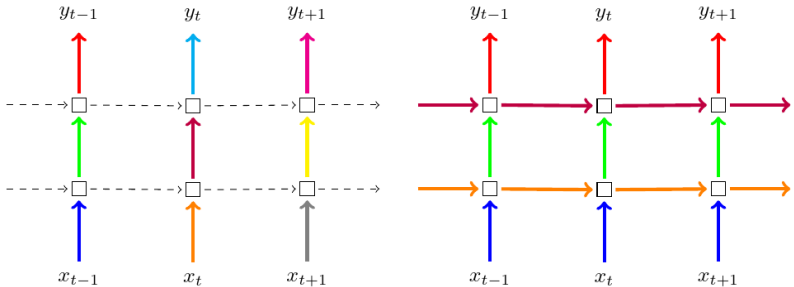
What to use for classification?

- The last hidden state?
- The sum or the average of each element states?
- Max pooling?



Dropout in RNNs

- Dropout along memory updates can inhibit learning of effective gating;
- only apply dropout 'vertically'; or fix random mask: 'variational RNN'



(a) Naive dropout RNN (b) Variational RNN

[Gal and Ghahramani, 2016]

References I



Gal, Y. and Ghahramani, Z. (2016).
A theoretically grounded application of dropout in recurrent neural networks.
Advances in neural information processing systems, 29.