From Seq2Seq to Chord2Vec

Sephora Madjiheurem

July 5, 2016

1 Sequence-to-sequence: Recap

Sequence-to-sequence models allow to learn a mapping of input sequences of varying lengths to output sequences also of varying lengths [1]. It uses a neural network known as RNN Encoder-Decoder. A LSTM (encoder) is used to map the input sequence to a fixed length vector, and another LSTM (decoder) is then used to extract the output sequence from this vector. The general goal is to estimate $p(y_1, \ldots, y_{T'}|x_1, \ldots, x_T)$, where $x_1, \ldots, x_T$ and $y_1, \ldots, y_{T'}$ are the input and output sequences respectively, and T' and T need not to be equal.

The objective is given by:

$$\max_{\theta} \frac{1}{|\mathcal{T}|} \sum_{(X,Y)\in\mathcal{T}} \log p(Y|X, \theta)\,, \tag{1}$$

where Y is a correct output given the input X and $\mathcal{T}$ is the training set and θ is the set of the model parameters. The encoder and decoder are jointly trained to maximize the objective according to θ.

The model estimates the conditional probability $p(y_1, \ldots, y_{T'}|x_1, \ldots, x_T)$ by first obtaining the fixed-length vector representation v of the input sequence (given by the last state of the LSTM encoder) and then computing the probability of $y_1, \ldots, y_{T'}$ with the LSTM decoder:

$$p(y_1, \ldots, y_{T'}|x_1, \ldots, x_T) = \prod_{t=1}^{T'} p(y_t|v, y_1, \ldots, y_{t-1}) \tag{2}$$

1.1 Attention mechanism

TODO

2 Chord2Vec

The seq2seq model can be used to learn an embedding of chords. In this setting, an input sequence is a sequence of notes in a chord (in some fixed ordering) and the output sequence are the sequences of notes in context chords. Thus, we jointly train the encoder and decoder to learn the context of a chord. Then, the last state of the LSTM encoder, that gives fixed-dimensional vector representations of input chords, is the desired chords embedding.

References

[1] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. *CoRR*, abs/1409.3215, 2014. URL `http://arxiv.org/abs/1409.3215`.