4. RNN Decoding
Decoding

3:44 / 3:44 1.0x

and thereby unravel a sentence using evolving state

from the RNN at each point predicting

what the next word is.

And once we have this architecture that

can generate sequences, such as sentences,

we can start any vector coming from a sentence, image,

or other context and unravel it as a sequence

as a natural language sentence.

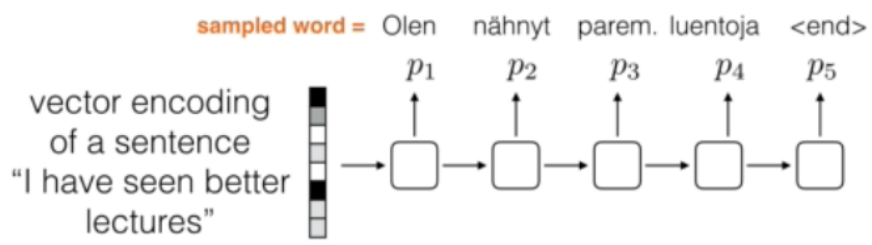
End of transcript. Skip to the start.

Video
Download video file

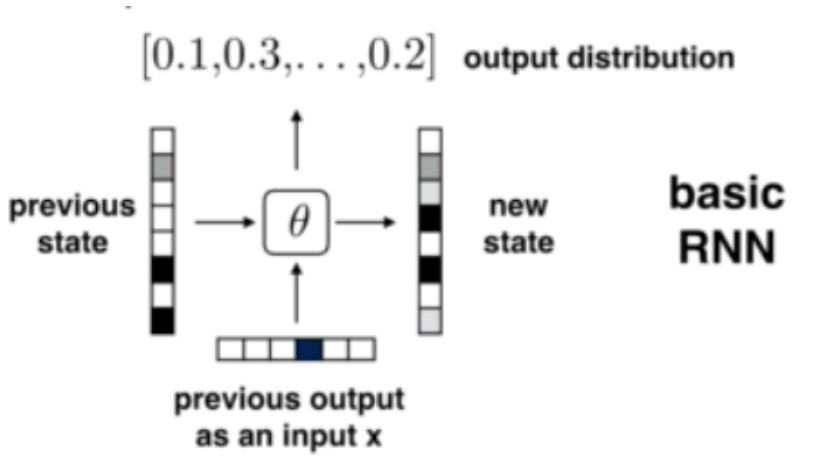
Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Decoding with RNN

1/1 point (graded)
Now, we would like to decode a feature vector with RNN's. The picture below illustrates how a vector encoding of the English sentence "I have seen better lectures" is translated into a sentence of a foreign language.



Unlike in encoding, at each step, an output distribution p_t is produced in a decoding RNN.



$$s_t = \tanh(W^{s,s}s_{t-1} + W^{s,x}x_t) \quad \text{state}$$
$$p_t = \text{softmax}(W^o s_t) \quad \text{output distribution}$$

Now, which of the following is true about decoding RNN's?
(Choose all those apply.)

- [ ] In the translating example above, the output probability distribution p_t is fed as an input to the next step
- [ ] The probability distribution p_t is the same at each step, just like how parameters are shared between steps
- [x] In the first image, the foreign word "Olen" in the above picture is a "sampled" result from a distribution the RNN produced. ✔

✔

Solution:

As shown in the figure, it is the previous output x_t but not the output probability ditribution p_t is fed into the next step. The probability distribution is different at each step as it propogated from the begining state. With the probability distribution at each step, the output word is then sampled from the distribution.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Predictions

1/1 point (graded)
Suppose we are building an RNN model to translate images into sentences, as described in the lecture. Which of the following is only done in generating the predictions of sentences from a trained RNN given the test images but not in the training process?

- (•) Feeding the sampled output as part of the input to the next time step ✔
- ( ) Calculating what percentage of words the RNN correctly generated
- ( ) Feeding in the hidden state as input each time step

Solution:

All of the above are done during training except using the sampled output as input. In training, you use the true words specified as input for the next time step. However, in testing you want to predict the sentence on your own, so you use the sampled output at one time step as the input for the next step.

Submit You have used 2 of 2 attempts

Answers are displayed within the problem