Recurrent Neural Networks

Sampling novel sequences

deeplearning.ai

After you train a sequence model, one of the ways you can informally get a sense of what is learned is to have a sample novel sequences.

Sampling a sequence from a trained RNN

$P(y^{<1>}, \ldots, y^{<T_x>})$

Training:

Sapling:

$P(a) P(aaron) \ldots P(zulu) P(<UNK>) \longrightarrow$ np.random.choice

$P(\_ | the)$

$<EOS>$

$<UNK>$

Andrew Ng

Character-level language model

advantage

disadvantage: ① not as good as word-level language models at capturing long range dependancies between how the earlier parts of the sentence also affect the later parts of the sentence.

→ Vocabulary = [a, aaron, …, zulu, <UNK>] ←

→ Vocabulary = [a, b, c, … , z, ⎵, ., ,, ;, 0, …, 9, A, …, Z]

② computationally expensive $y^{<1>} y^{<2>} y^{<3>} y^{<4>}$
to train.

③ end up with much more, much longer sequences.

Cat   average
↑ ↑ ↑ ↑ …
$y^{<1>} y^{<2>} y^{<3>} y^{<4>}$

Mau

$\hat{y}^{<1>}$ $\hat{y}^{<2>}$ $\hat{y}^{<3>}$ $\hat{y}^{<T_y>}$

$a^{<0>} \rightarrow$ $\boxed{a^{<1>}} \rightarrow \boxed{a^{<2>}} \rightarrow \boxed{a^{<3>}} \rightarrow \quad \cdots \quad \rightarrow \boxed{a^{<T_y>}}$

$x^{<1>}$ $\hat{y}^{<1>}$ $\hat{y}^{<2>}$ $\hat{y}^{<T_x-1>}$

Andrew Ng
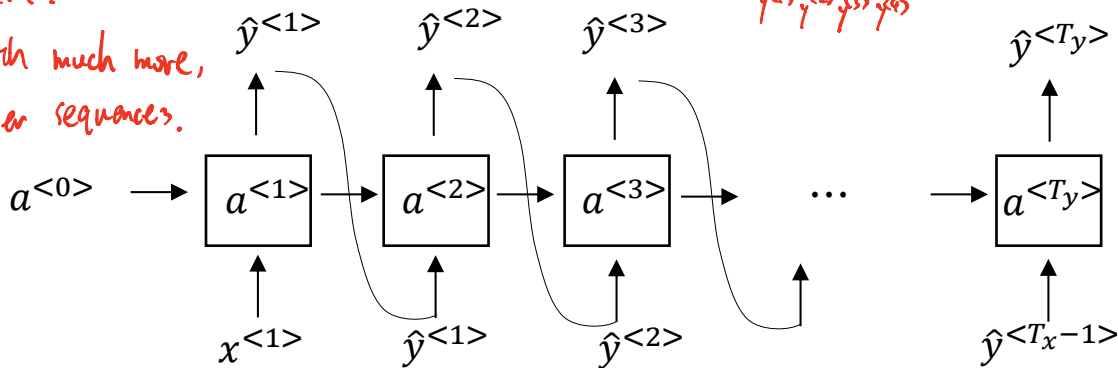
Sequence generation

News

President enrique peña nieto, announced sench's sulk former coming football langston paring.

"I was not at all surprised," said hich langston.

"Concussion epidemic", to be examined. ←

The gray football the told some and this has on the uefa icon, should money as.

Shakespeare

The mortal moon hath her eclipse in love.

And subject of this thou art another this fold.

When besser be my love to me see sabl's.

For whose are ruse of mine eyes heaves.