

Course 5: Sequence models

Content:

Week 1: Recurrent Neural Networks

Week 2: Natural Language Processing

• Word Embeddings

Week 3: Sequence models • Attention
Mechanism

Coursera

Deep Learning Specialization



Sequence data: Speech recognition, music generation, etc.

Recurrent Neural network

each step, rnn passes on as activation to next step

for it to use

scan data left \rightarrow right

parameters it uses for each time step are shared.

* limitation \rightarrow Pred. at certain time uses inputs or
uses info. from inputs earlier in the sequence but not
info. later in seq.

Types of RNN:

$1 \rightarrow 1$

$1 \rightarrow \text{many}$

$\text{many} \rightarrow 1$

$\text{many} \rightarrow \text{many}$

Problems RNN

⇒ Vanishing gradients

→ not good at capturing very long term
dependencies

↓
Gated Recurrent Unit (GRU)

↓
simpler
model

Long short Term memory (LSTM)

↓
more effective

(gates)

target T_f , - update T_u , output P_o .

(state)

cell $c^{(t)}$, hidden $a^{(t)}$

candidate value $\tilde{c}^{(t)}$

Translation system:

(beam search)