

NATURAL LANGUAGE PROCESSING

LECTURE 4: RNN

goorm

KAIST AI
Graduate School of AI

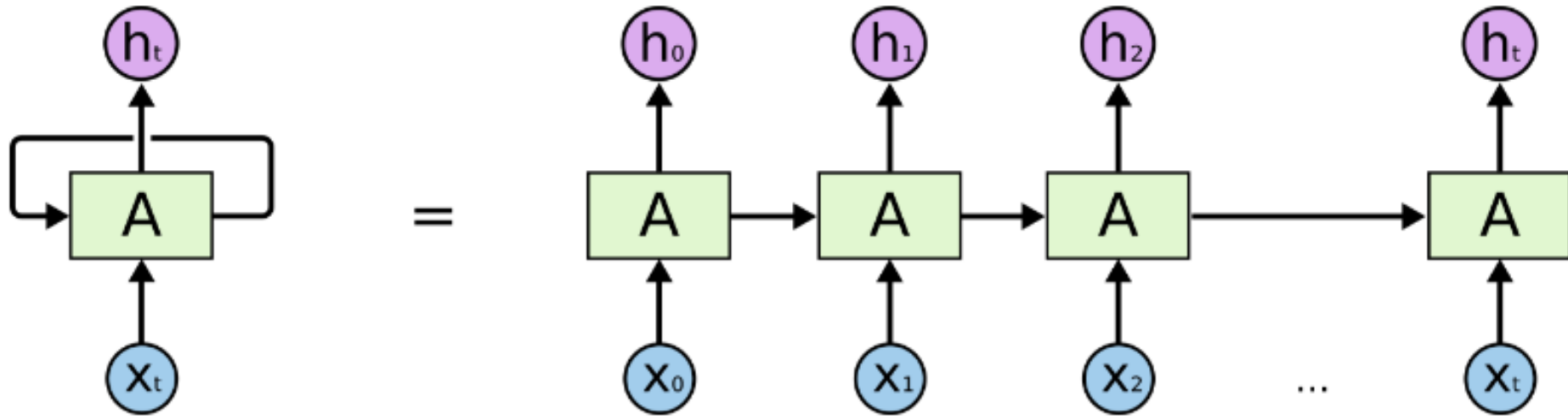


Index

- Character Language Model
- RNN
- Vanishing Gradient

RNN의 여러가지 형태

가장 기본적인 구조

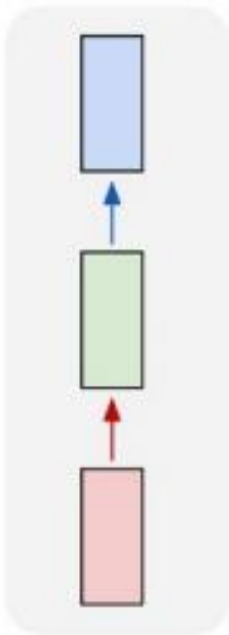


An unrolled recurrent neural network.

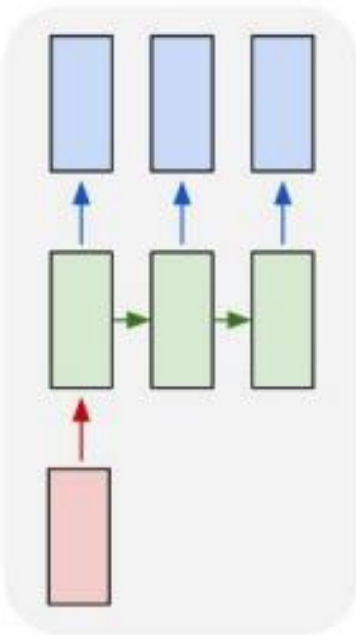
RNN의 여러가지 형태

기본 neural networks 구조

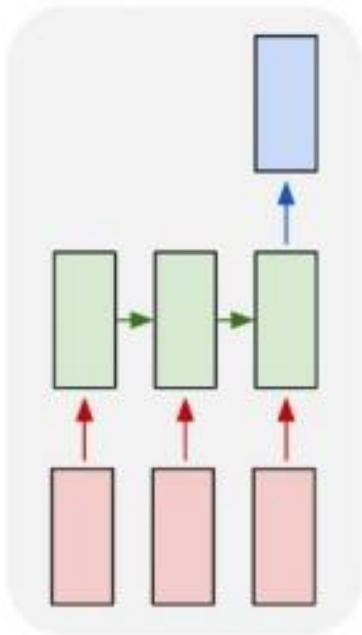
one to one



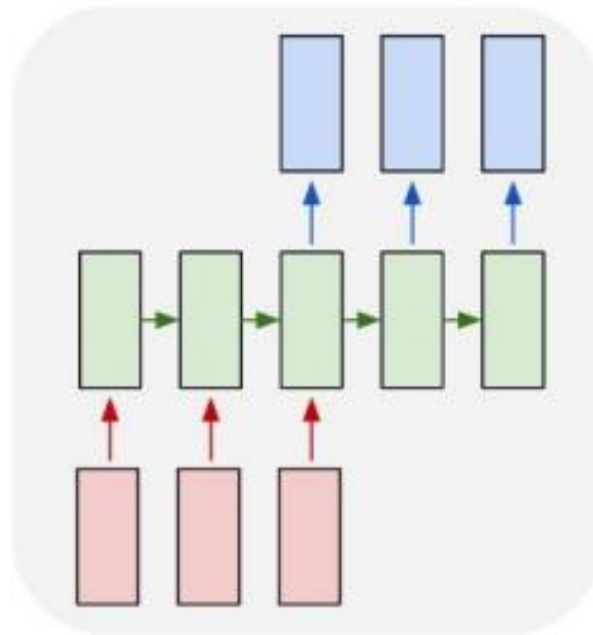
one to many



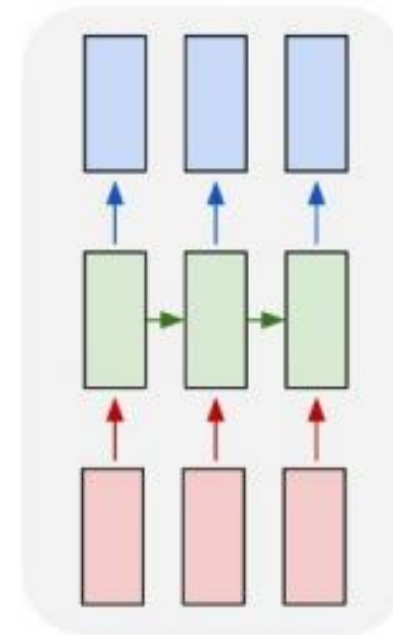
many to one



many to many



many to many



Vanilla Neural Networks

Character-level Language Model

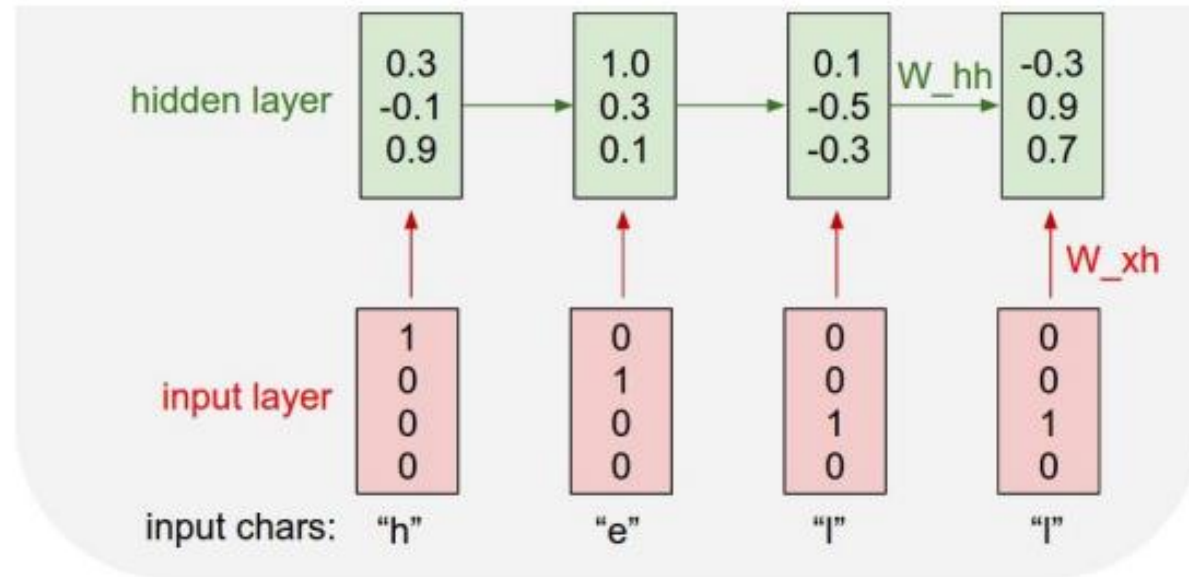
- Example of training sequence “hello”

Character-level language model example

Vocabulary:
[h,e,l,o]

Example training
sequence:
“hello”

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$



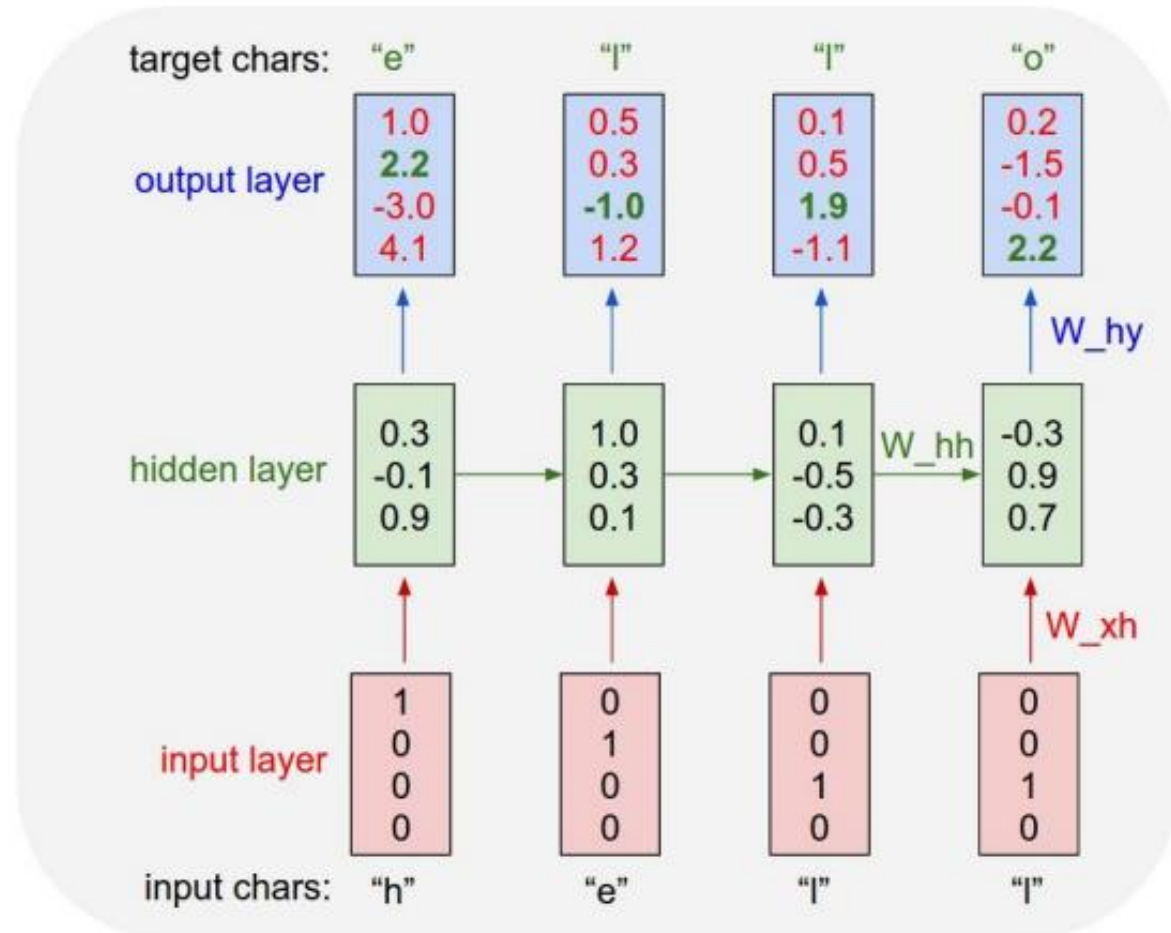
Character-level Language Model

- Example of training sequence “hello”

Character-level language model example

Vocabulary:
[h,e,l,o]

Example training sequence:
“hello”



Character-level Language Model

- Example of training sequence “hello”

**Example:
Character-level
Language Model
Sampling**

Vocabulary:
[h,e,l,o]

At test-time sample
characters one at a time,
feed back to model

