

Sequence Modeling

성균관대학교 소프트웨어학과
이 지 형

Sequential Data Modeling

- ▶ **Sequential Data**

- ▶ Most of data are sequential
- ▶ Speech, Text, Image, ...

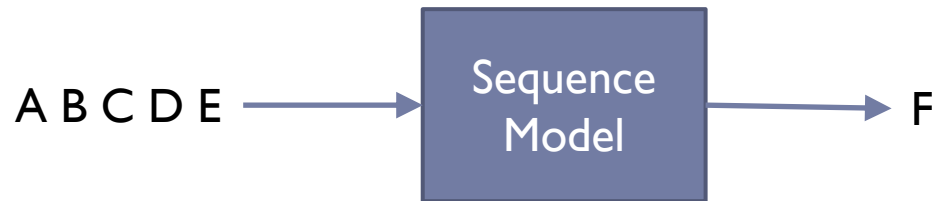
- ▶ **Deep Learnings for Sequential Data**

- ▶ **Convolutional Neural Networks (CNN)**
 - ▶ Try to find local features from a sequence
- ▶ **Recurrent Neural Networks: LSTM, GRU**
 - ▶ Try to capture the feature of the past

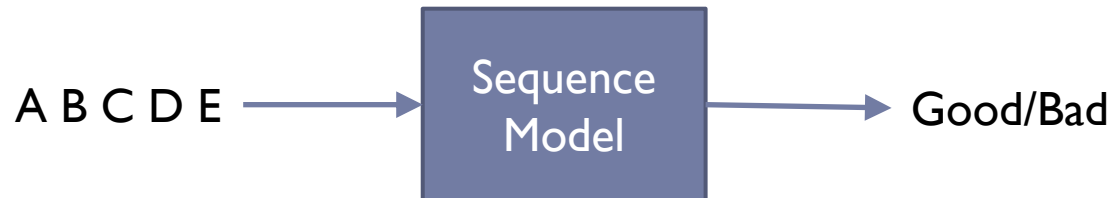
Sequential Data Modeling

▶ Three Types of Problems

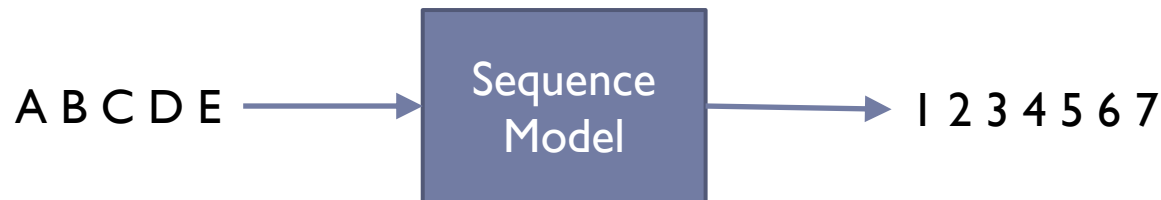
▶ Next Step Prediction



▶ Classification



▶ Sequence Generation



Sequential Data Modeling

▶ Sequence Generation


▶ Machine Translation

This is a very good wine → C'est un très bon vin

▶ Speech Recognition

 → This is a very good wine

▶ Image Caption Generation

 → A bird is flying

Types of Processes

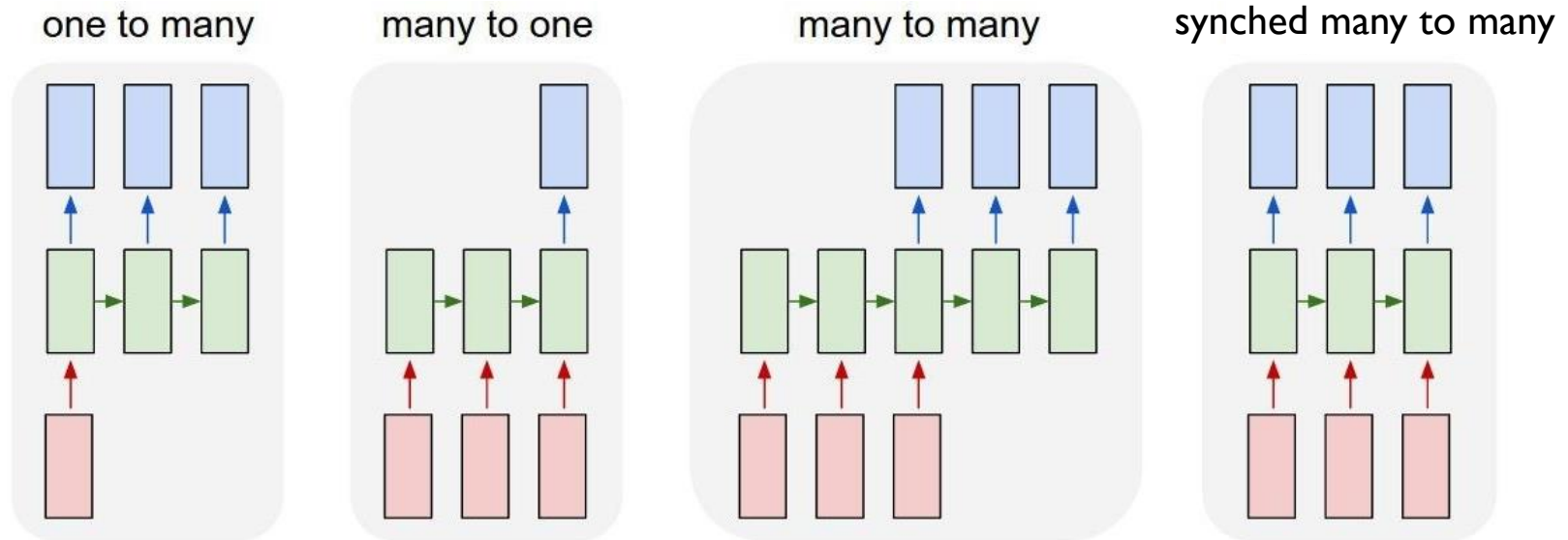
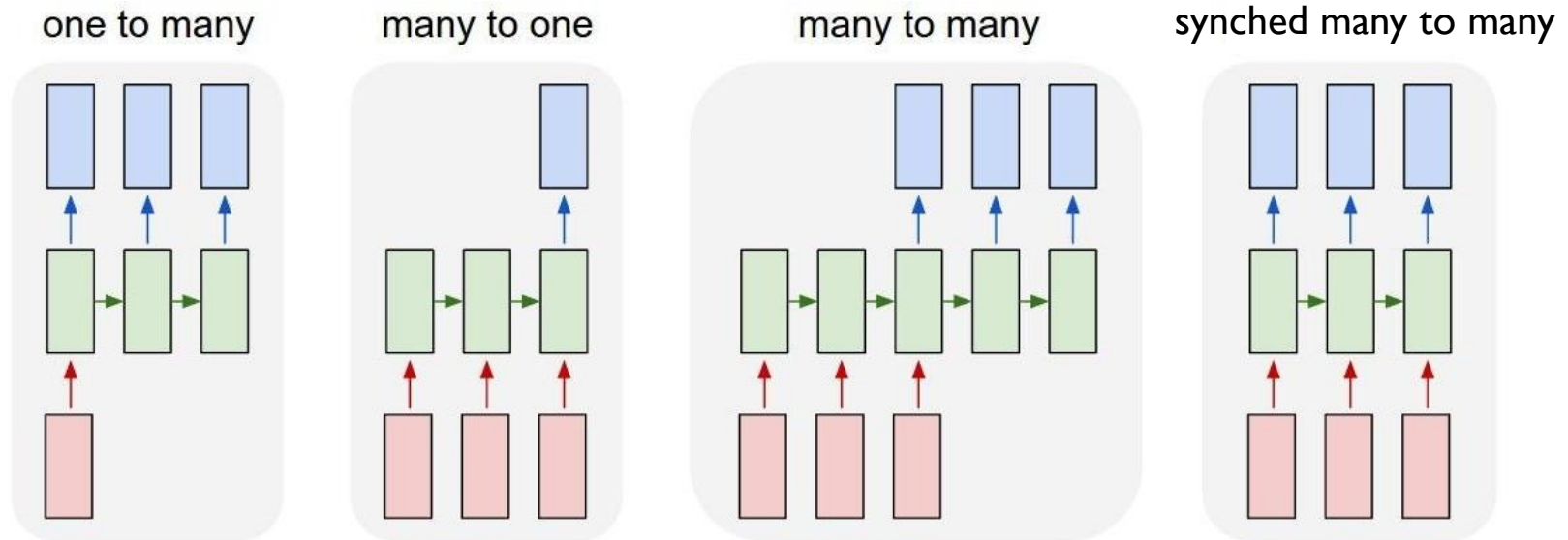


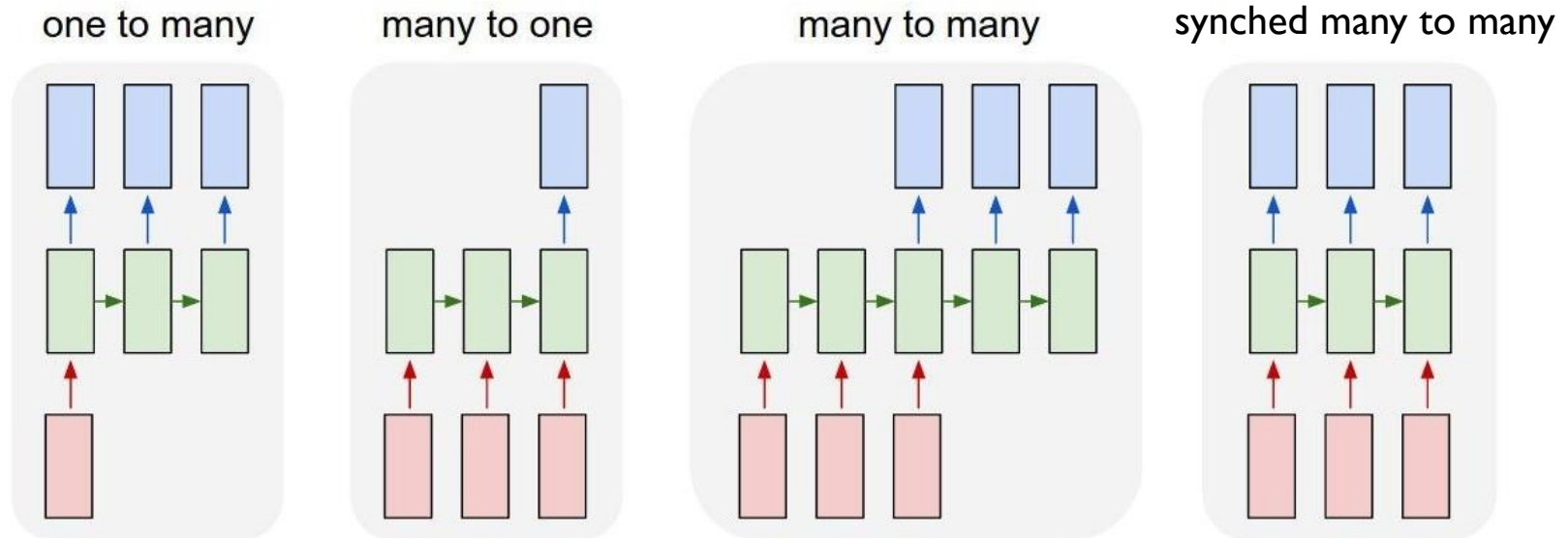
Image Captioning
Image → sequence of words

Types of Processes



→ Sentiment Classification
sentence → sentiment

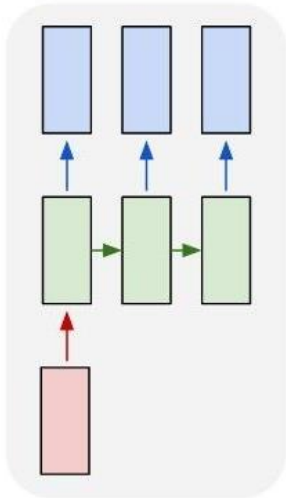
Types of Processes



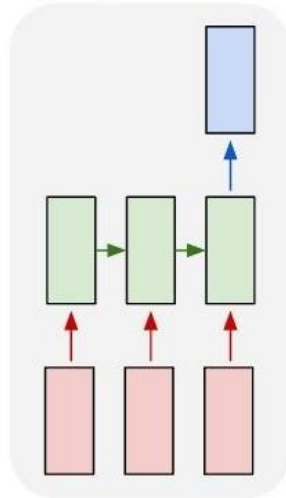
Machine Translation
sentence → sentence

Types of Processes

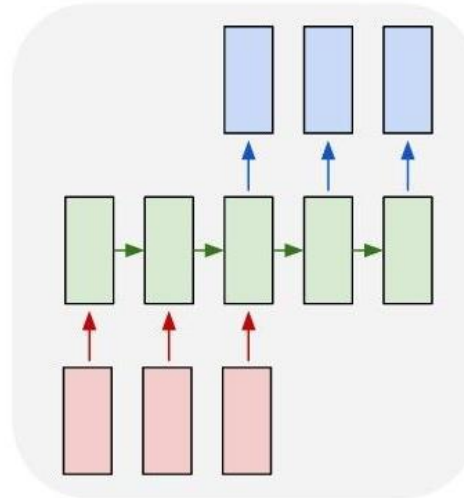
one to many



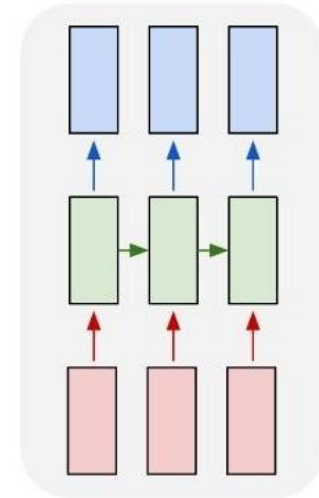
many to one



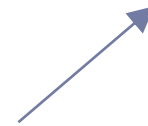
many to many



synched many to many



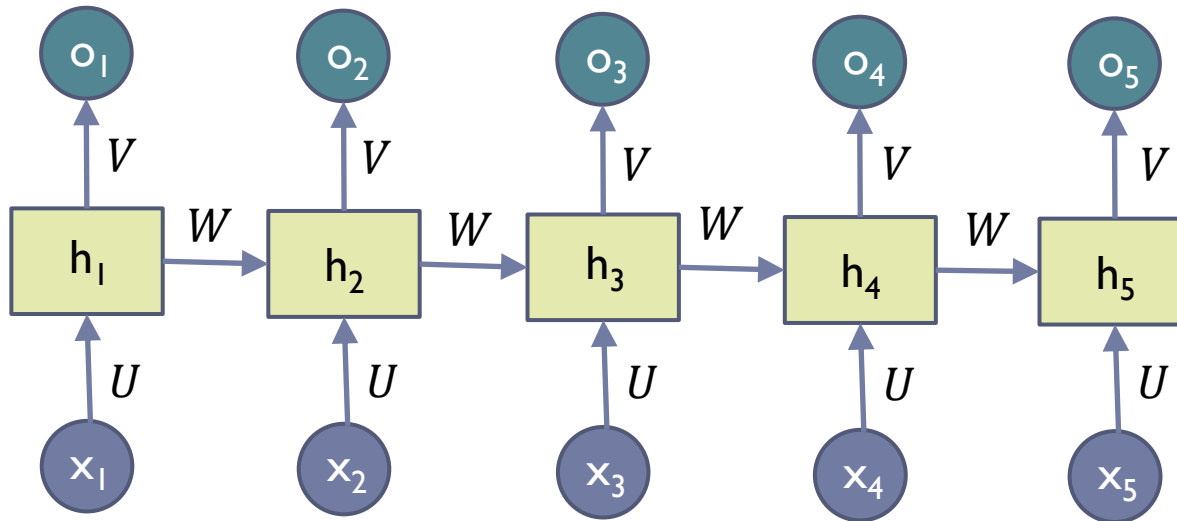
Stock Price Prediction,
Prediction of next word



Synced Many to Many

▶ Training

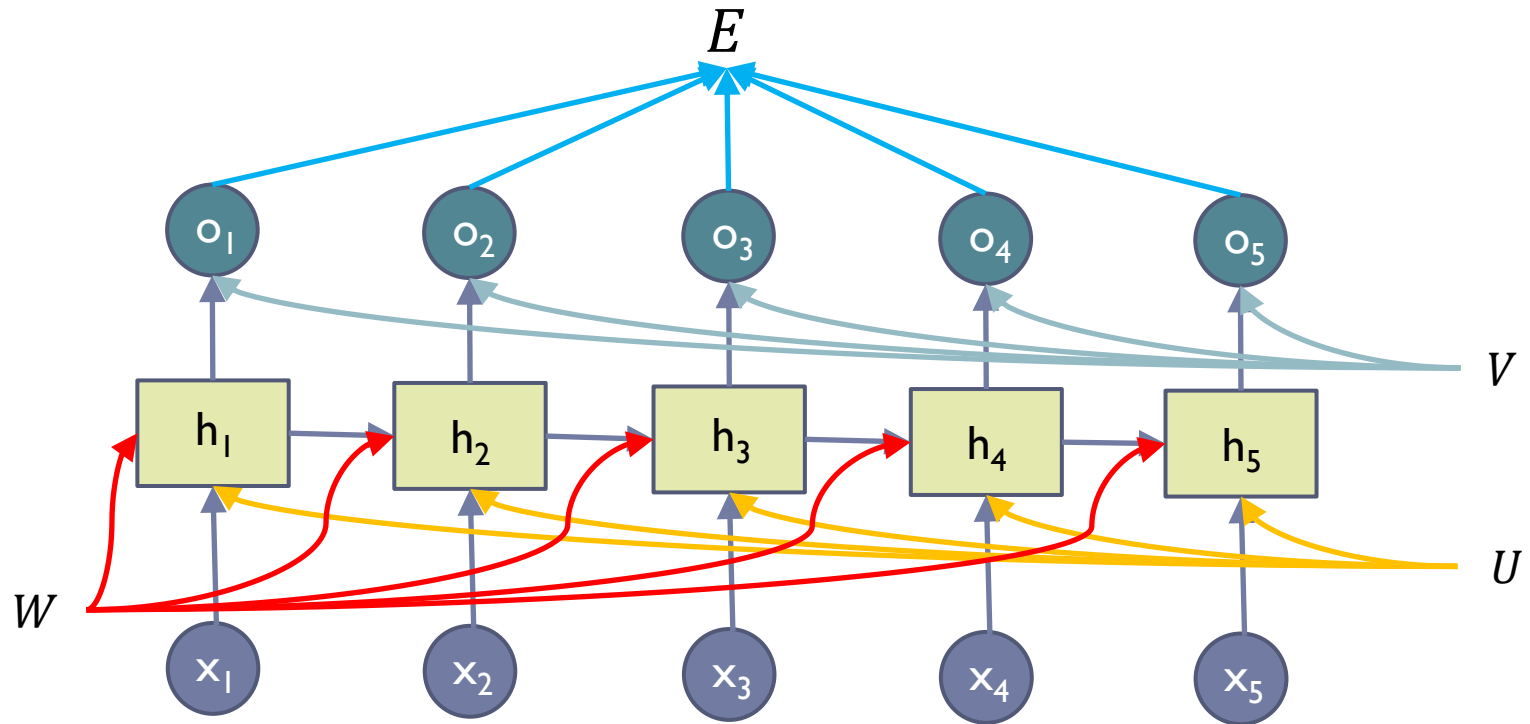
$$x_1 x_2 x_3 \cdots x_n \rightarrow y_1 y_2 y_3 \cdots y_n$$



$$E = \sum_{i=1}^n (y_i - o_i)^2$$

Synced Many to Many

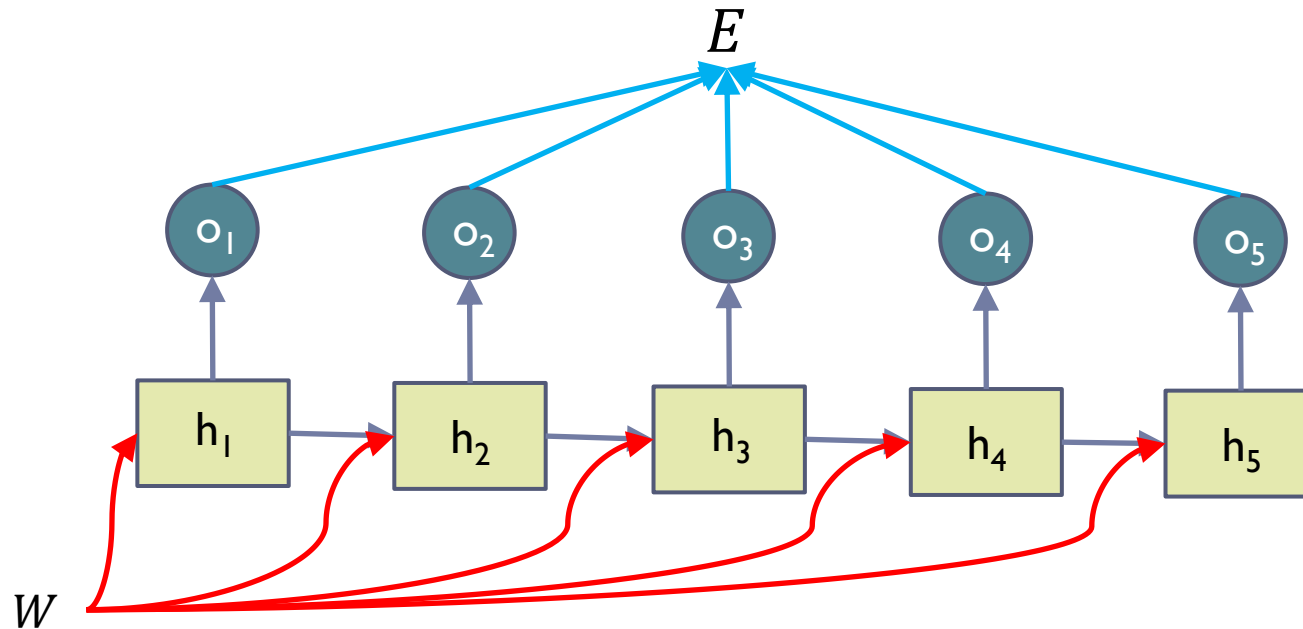
► Training



$$\frac{\partial E}{\partial w} = ?$$

Synced Many to Many

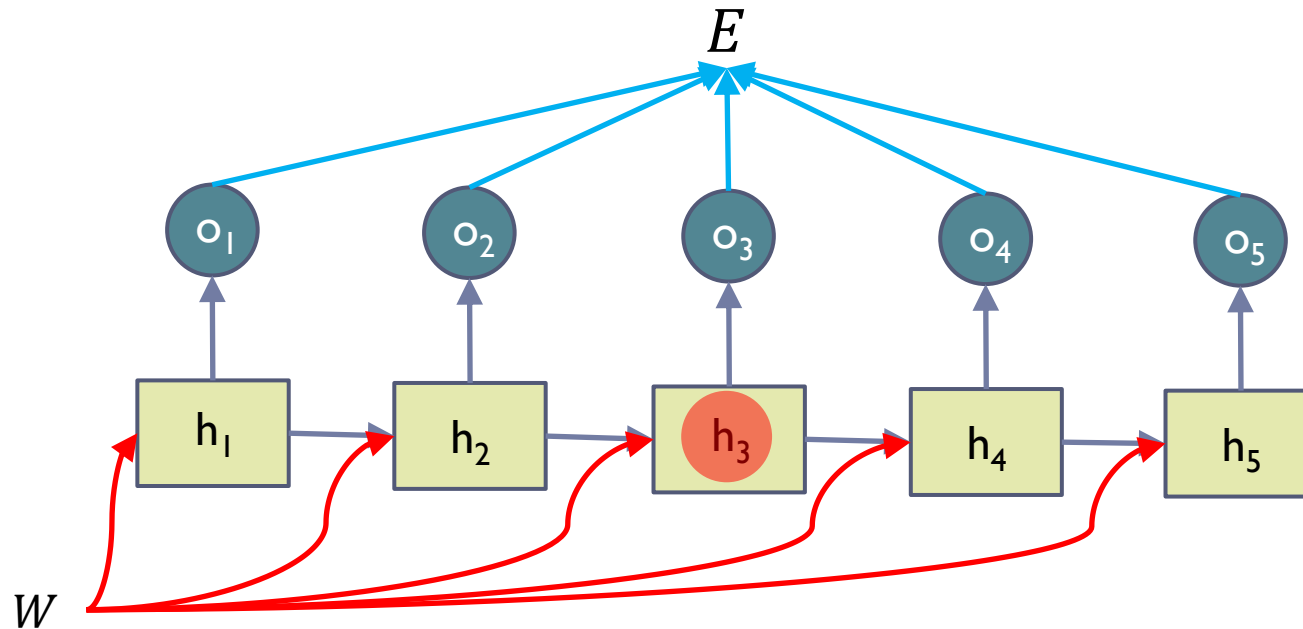
► Training



$$\frac{\partial E}{\partial w} = \sum_{i=1}^n \frac{\partial E}{\partial h_i} \frac{\partial h_i}{\partial w}$$

Synced Many to Many

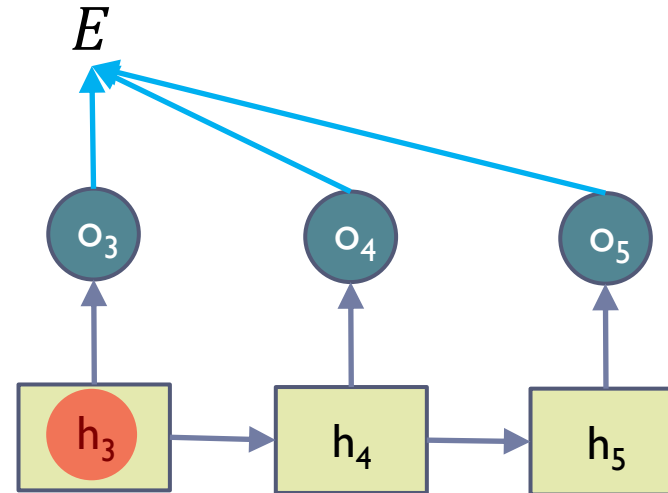
► Training



$$\frac{\partial E}{\partial w} = \sum_{i=1}^n \frac{\partial E}{\partial h_i} \frac{\partial h_i}{\partial w}$$

Synced Many to Many

► Training



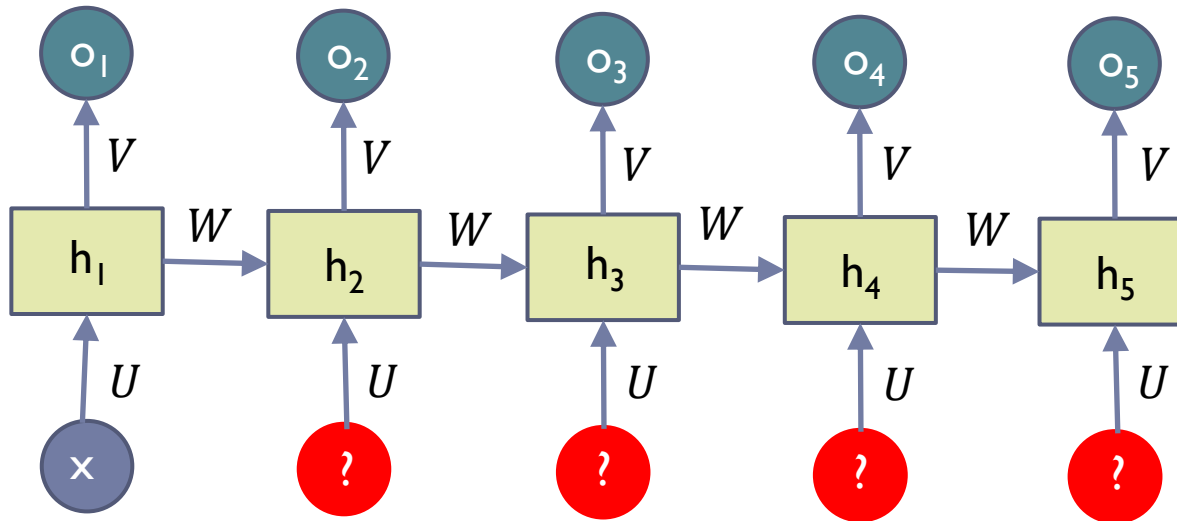
$$\frac{\partial E}{\partial h_i} = \frac{\partial E}{\partial o_i} \frac{\partial o_i}{\partial h_i} + \frac{\partial E}{\partial h_{i+1}} \frac{\partial h_{i+1}}{\partial h_i}$$

$$\frac{\partial E}{\partial w} = \sum_{i=1}^n \frac{\partial E}{\partial h_i} \frac{\partial h_i}{\partial w}$$

One to Many

▶ Training

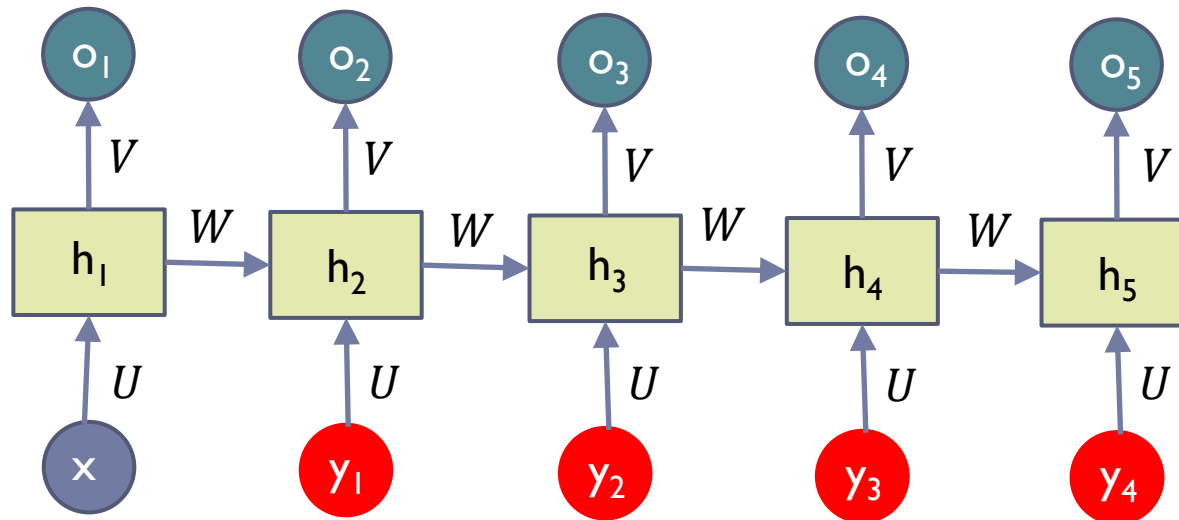
$$x \rightarrow y_1 y_2 y_3 \cdots y_n$$



One to Many

▶ Training

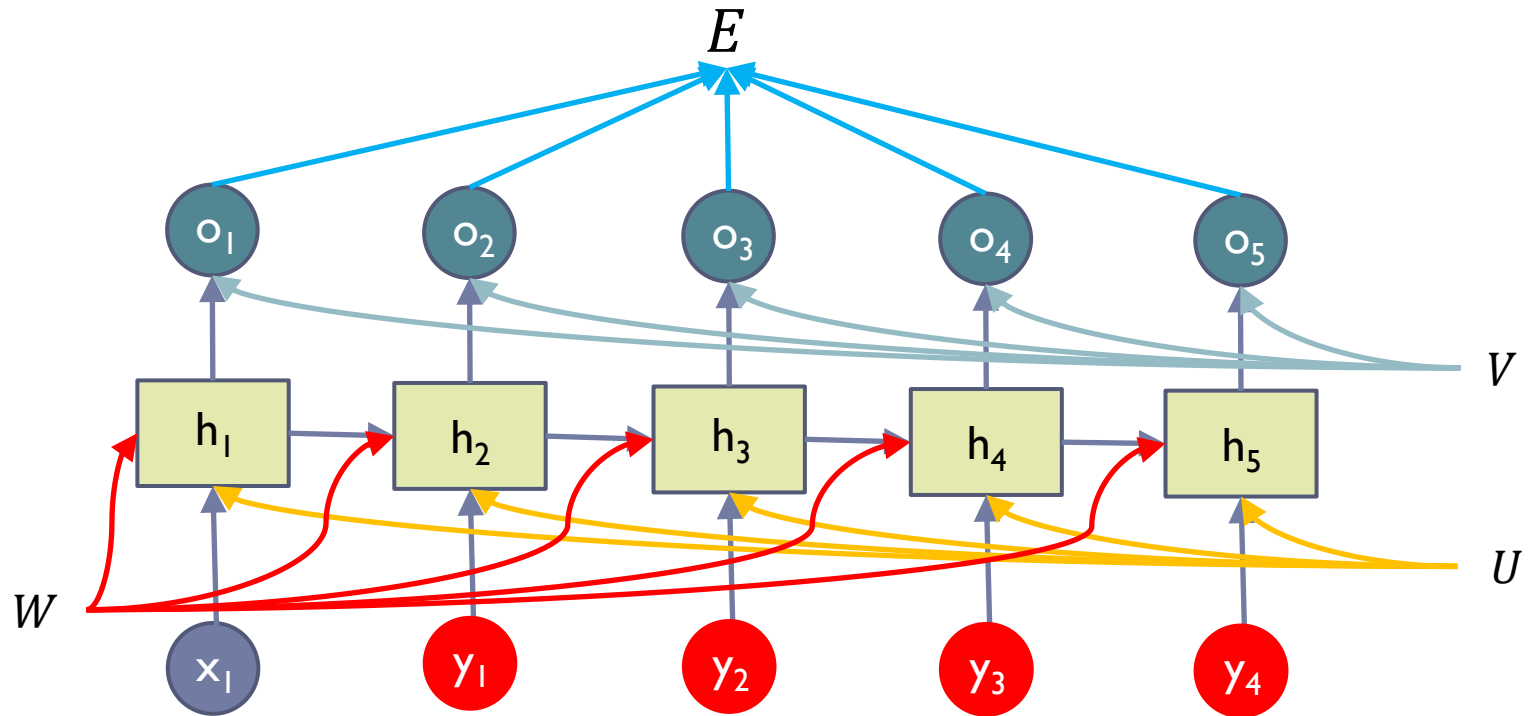
$$x \rightarrow y_1 y_2 y_3 \cdots y_n$$



$$E = \sum_{i=1}^n (y_i - o_i)^2$$

One to Many

► Training



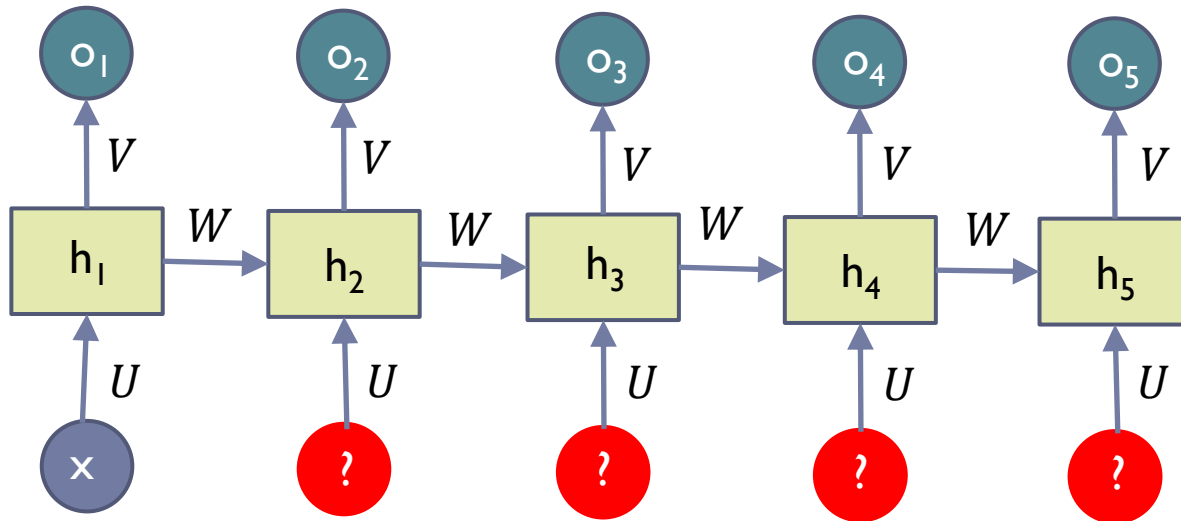
$$\frac{\partial E}{\partial w} = \sum_{i=1}^n \frac{\partial E}{\partial h_i} \frac{\partial h_i}{\partial w}$$

$$\frac{\partial E}{\partial h_i} = \frac{\partial E}{\partial o_i} \frac{\partial o_i}{\partial h_i} + \frac{\partial E}{\partial h_{i+1}} \frac{\partial h_{i+1}}{\partial h_i}$$

One to Many

▶ Testing

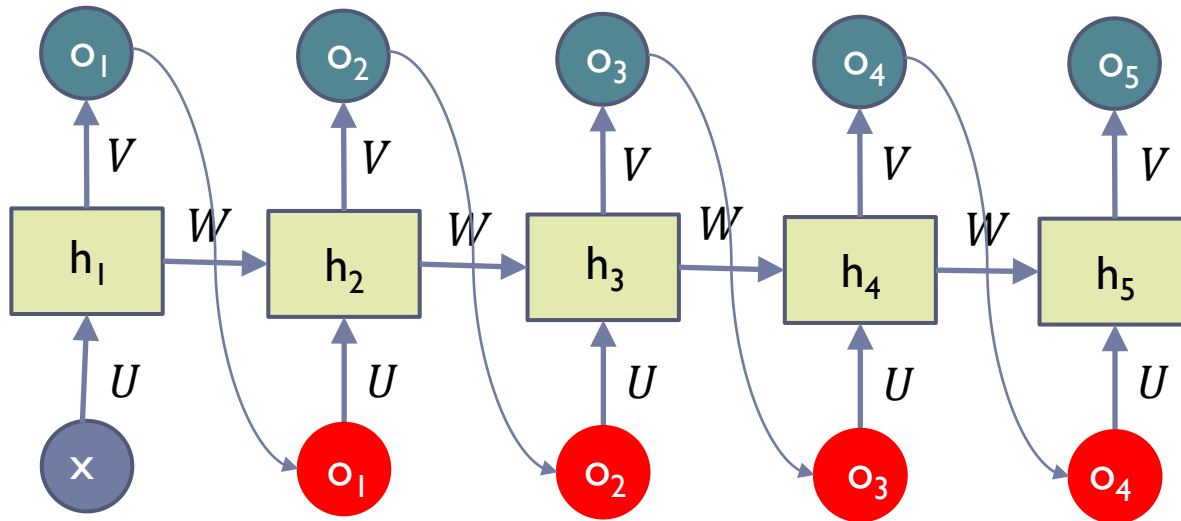
$x \rightarrow ???????$



One to Many

▶ Testing

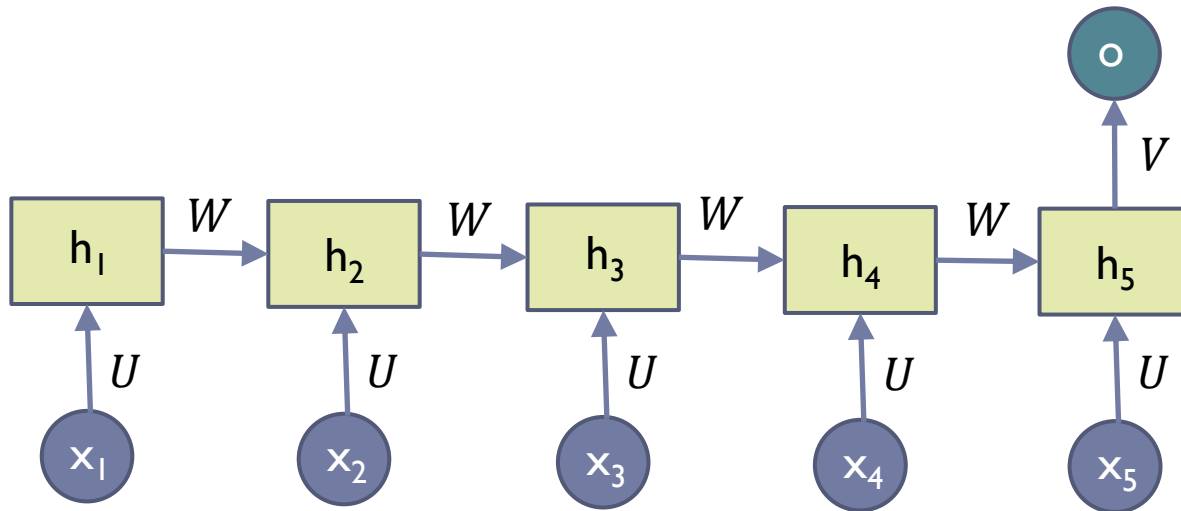
$x \rightarrow ???????$



Many to One

▶ Training

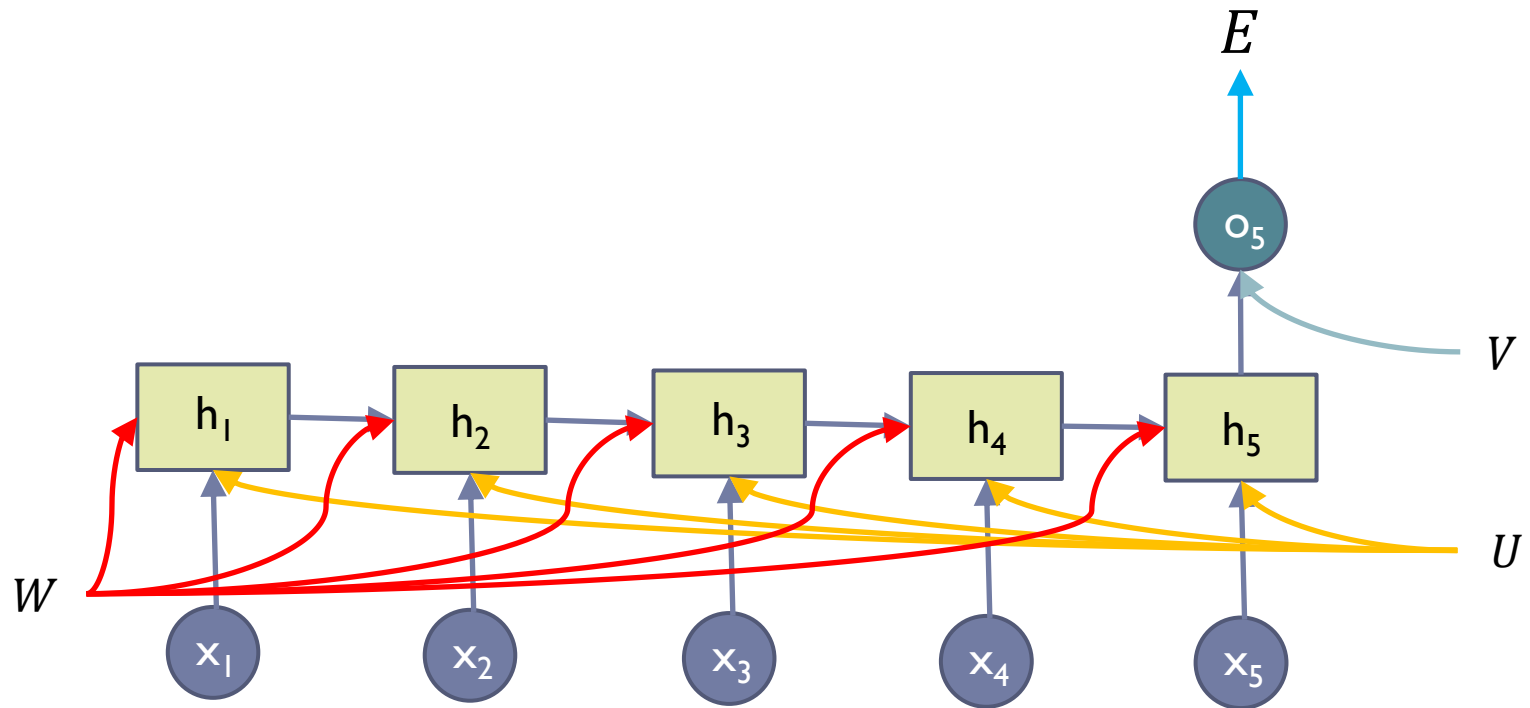
$$x_1 x_2 x_3 \cdots x_n \rightarrow y$$



$$E = (y - o)^2$$

Many to One

▶ Training



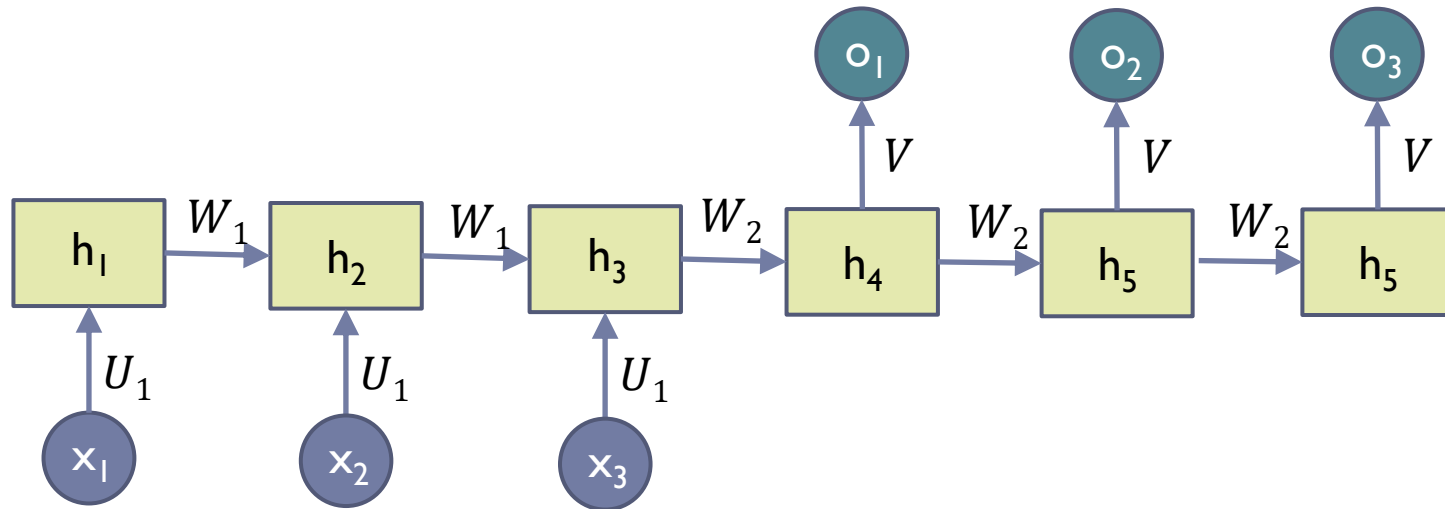
$$\frac{\partial E}{\partial w} = \sum_{i=1}^n \frac{\partial E}{\partial h_i} \frac{\partial h_i}{\partial w}$$

$$\frac{\partial E}{\partial h_i} = \frac{\partial E}{\partial h_{i+1}} \frac{\partial h_{i+1}}{\partial h_i}$$

Many to Many

▶ Training

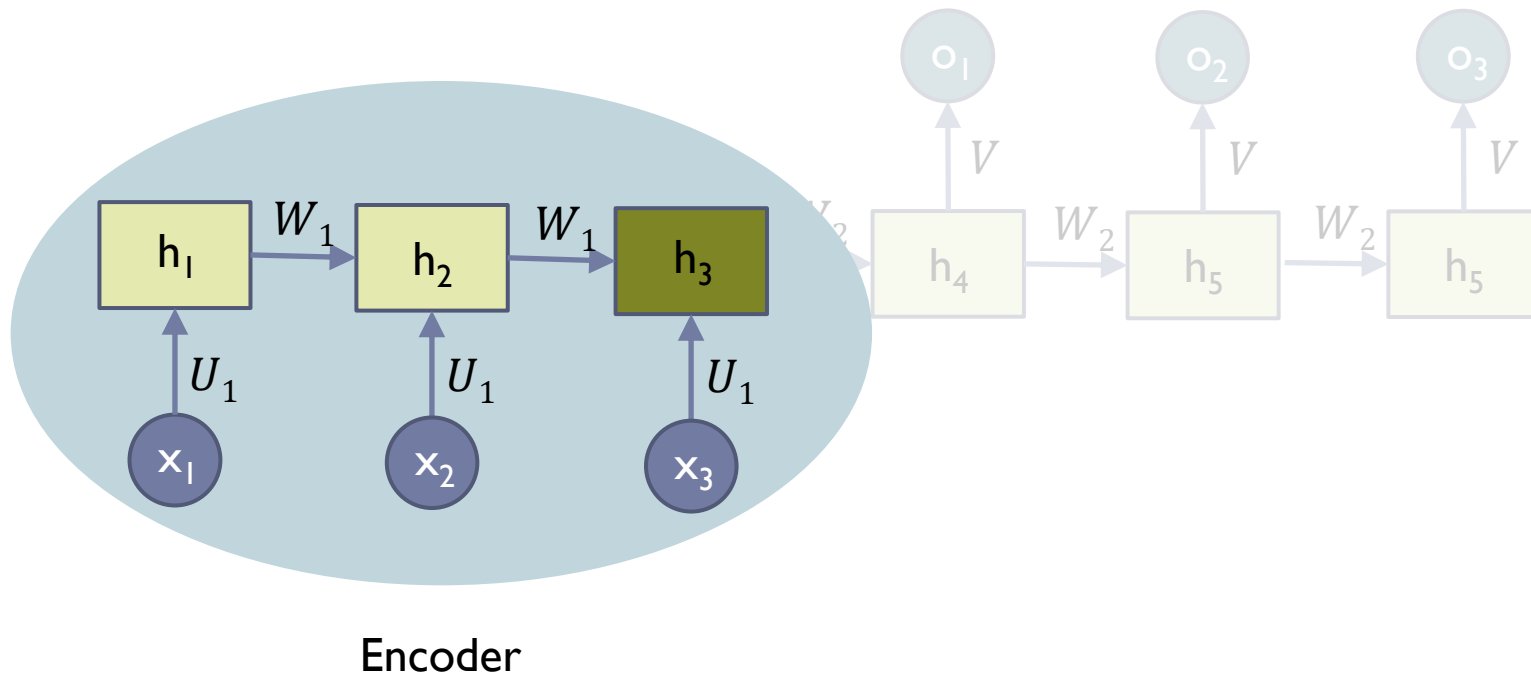
$$x_1 x_2 x_3 \cdots x_n \rightarrow y_1 y_2 y_3 \cdots y_n$$



Many to Many

▶ Training

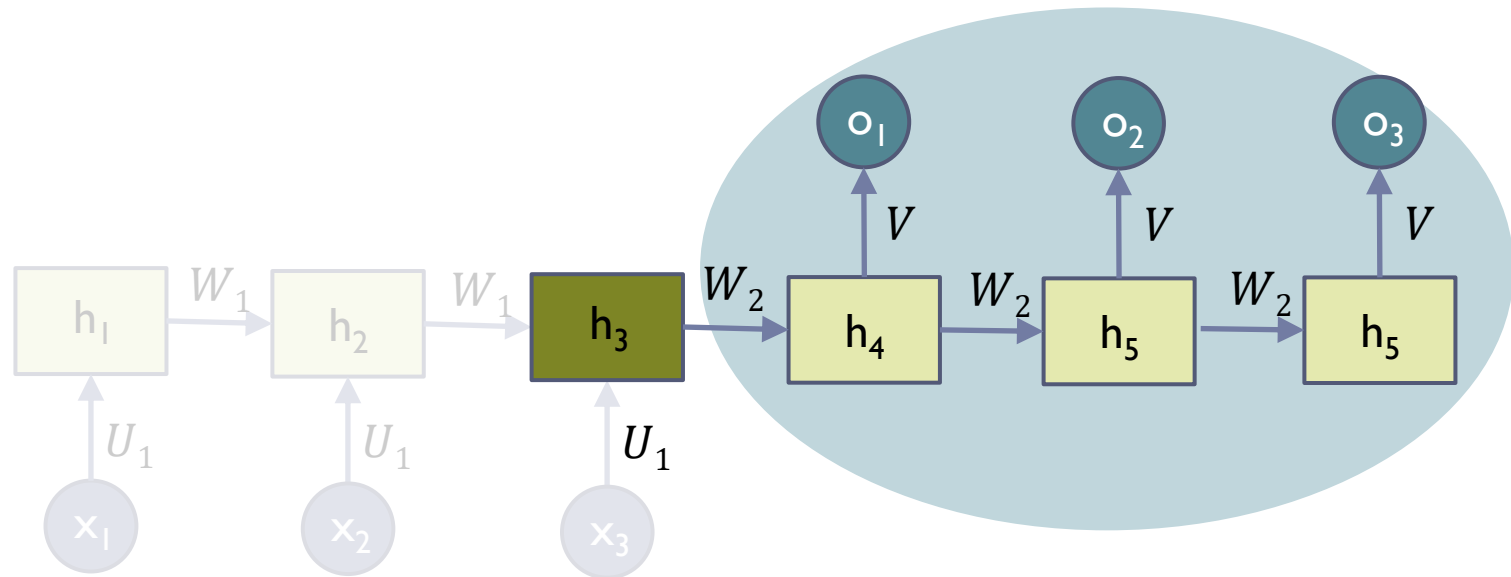
$$x_1 x_2 x_3 \cdots x_n \rightarrow y_1 y_2 y_3 \cdots y_n$$



Many to Many

▶ Training

$$x_1 x_2 x_3 \cdots x_n \rightarrow y_1 y_2 y_3 \cdots y_n$$



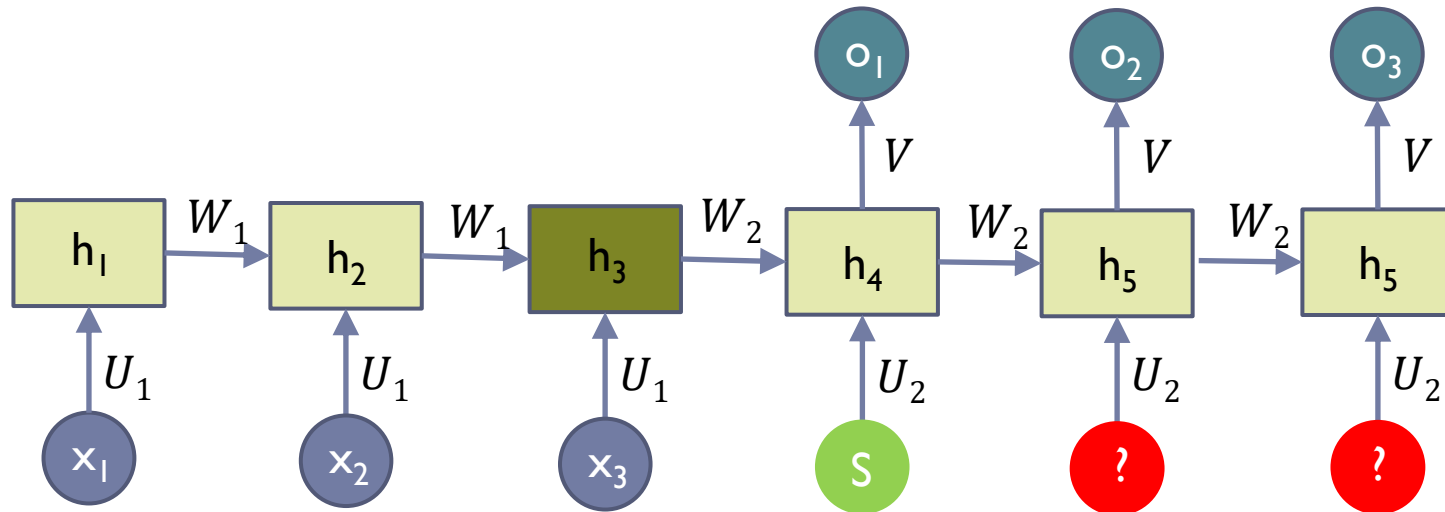
Decoder

Encoder

Many to Many

► Training

$$x_1 x_2 x_3 \cdots x_n \rightarrow y_1 y_2 y_3 \cdots y_n$$

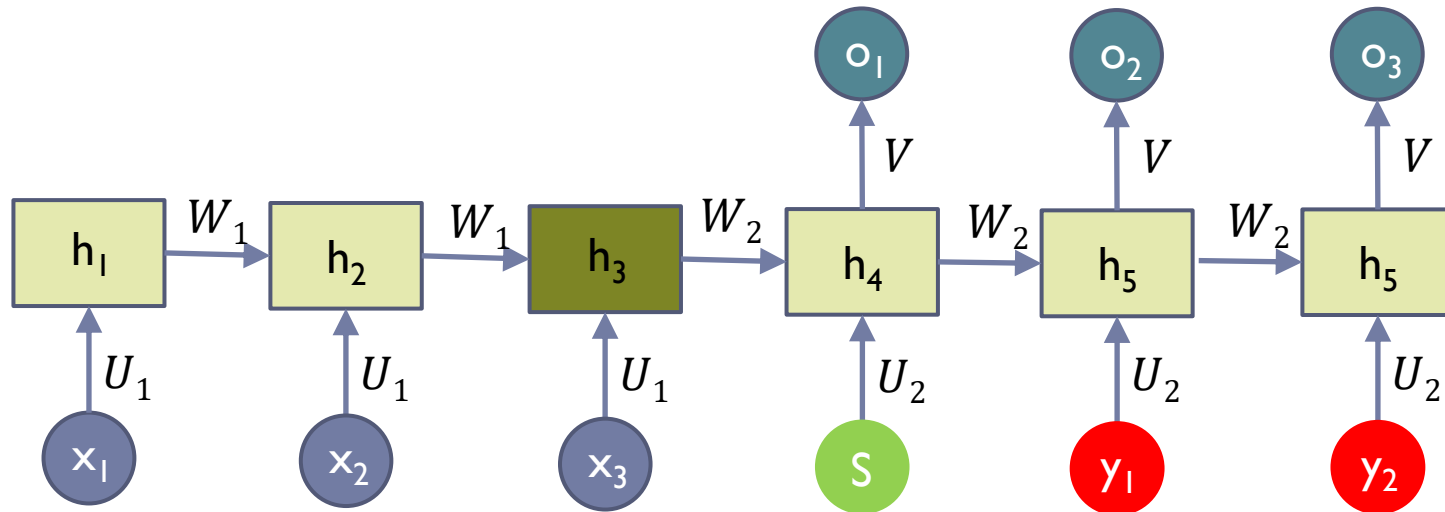


$$E = \sum_{i=1}^n (y_i - o_i)^2$$

Many to Many

► Training

$$x_1 x_2 x_3 \cdots x_n \rightarrow y_1 y_2 y_3 \cdots y_n$$



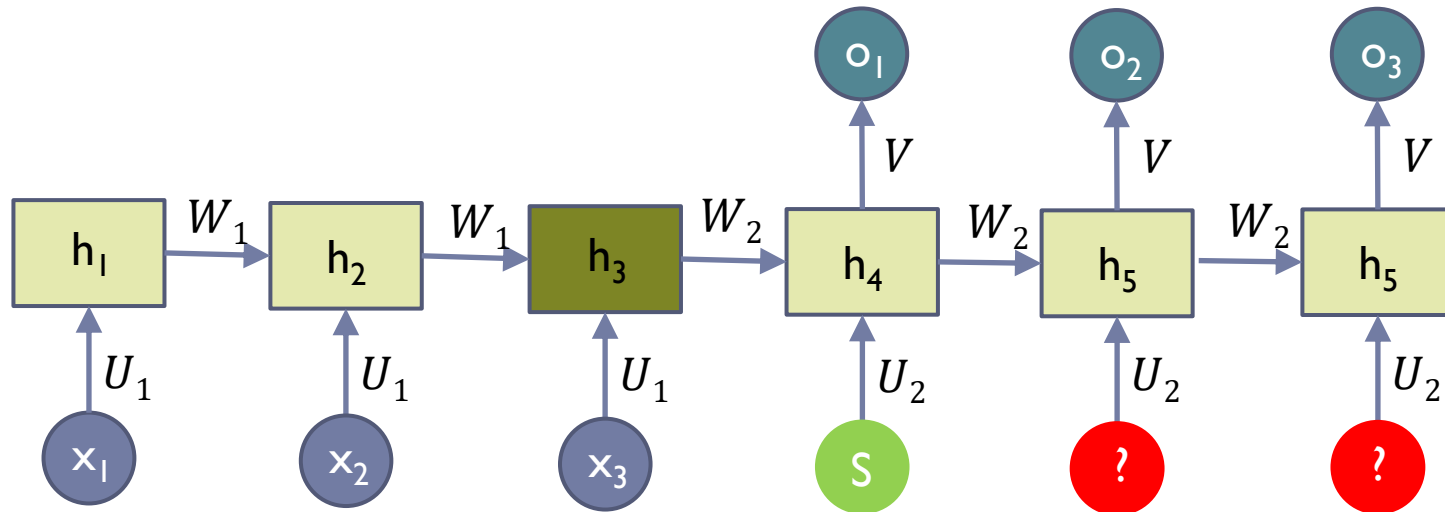
$$\frac{\partial E}{\partial w} = ??$$

Combination of [Many to One] and [One to Many]

Many to Many

▶ Testing

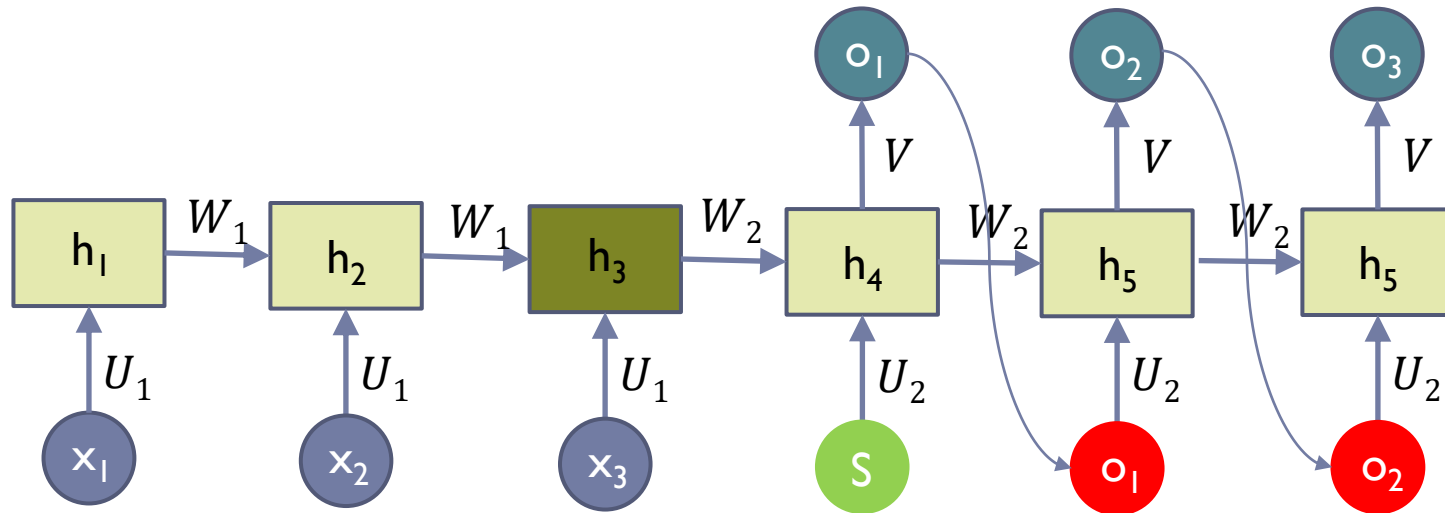
$$x_1 x_2 x_3 \cdots x_n \rightarrow ? ? ? ? ?$$



Many to Many

▶ Testing

$$x_1 x_2 x_3 \cdots x_n \rightarrow ? ? ? ? ?$$

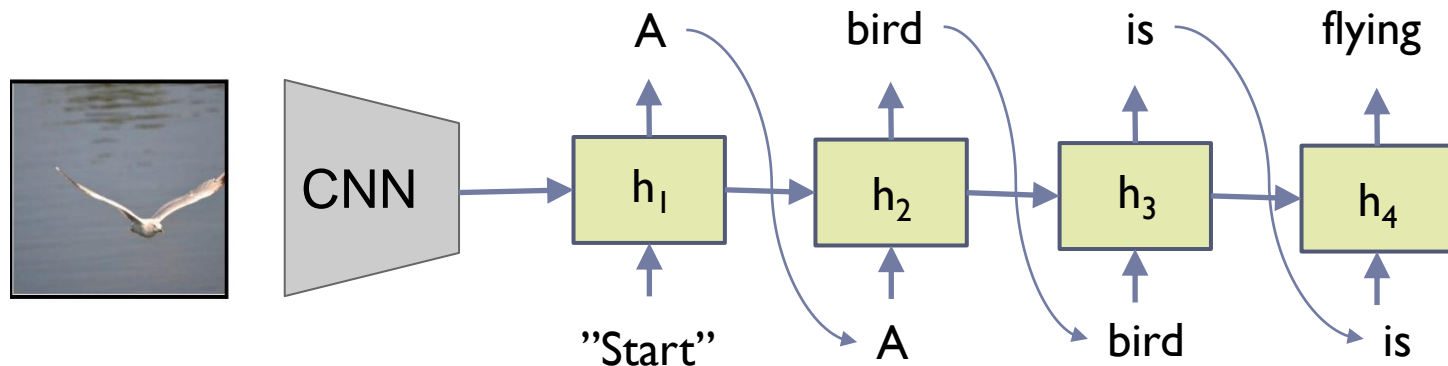


Example

▶ One to Many

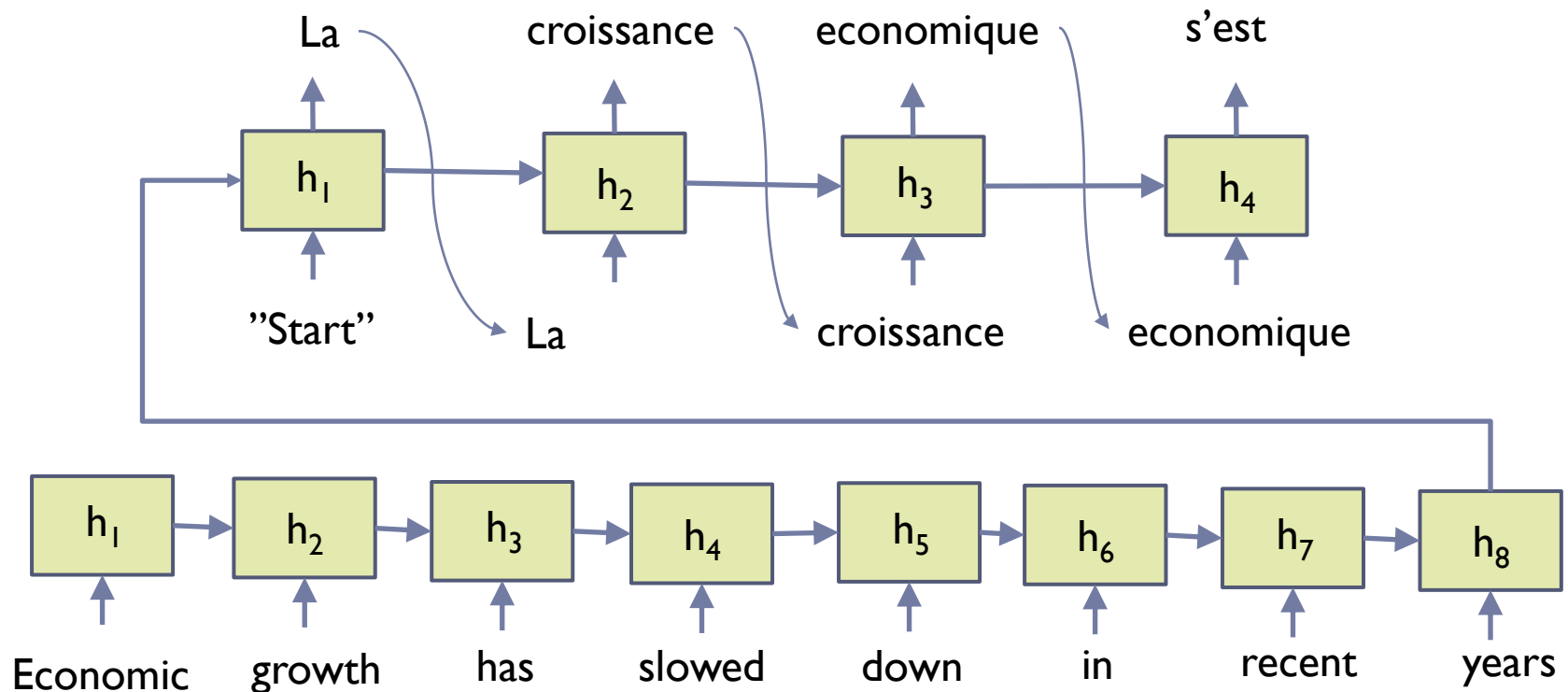
▶ Caption Generation

- ▶ Image is represented by a CNN
- ▶ Word Embedding at the input layer
- ▶ Softmax at the output layer



Example

- ▶ **Many to Many**
 - ▶ Word Embedding



Question and Answer