Idea Factory Intensive Program #2

딥러닝 홀로서기

#25

이론강의/PyTorch실습/코드리뷰

딥러닝(Deep Learning)에 관심이 있는 학생 발굴을 통한
딥러닝의 이론적 배경 강의 및 오픈소스 딥러닝 라이브러리 PyTorch를 활용한 실습

Acknowledgement

Sung Kim`s 모두를 위한 머신러닝/딥러닝 강의

- https://hunkim.github.io/ml/
- https://www.youtube.com/playlist?list=PLlMkM4tgfjnLSOjrEJN31gZATbcj_MpUm

Andrew Ng`s and other ML tutorials

- https://class.coursera.org/ml-003/lecture
- http://www.holehouse.org/mlclass/ (note)
- Deep Learning Tutorial
- Andrej Karpathy's Youtube channel

WooYeon Kim & SeongOk Ryu`s KAIST CH485 Artificial Intelligence and Chemistry

- https://github.com/SeongokRyu/CH485---Artificial-Intelligence-and-Chemistry

SungJu Hwang`s KAIST CS492 Deep Learning Course Material

Many insightful articles, blog posts and Youtube channels

Facebook community

- Tensorflow KR (https://www.facebook.com/groups/TensorFlowKR/)
- Pytorch KR (https://www.facebook.com/groups/PyTorchKR/)

Medium Channel and Writers

- Toward Data Science (https://towardsdatascience.com/)

Today`s Time Schedule

Assignment #5 Review	20 mins
Recurrent Neural Network	1 hour
Implement Basic RNN in Pytorch	1.5 hour

Dealing with Sequential Data

Imagine Your Boss Wants You to Solve a Problem Like…

Automatically generate caption with the given image

Predict whether a company would be bankrupted

Translate one sentence into another language

Classify whether the word is owns` name or not

Imagine Your Boss Wants You to Solve a Problem Like...

Automatically generate caption with the given image

Predict whether a company would be bankrupted

Translate one sentence into another language

Classify whether the word is owns` name or not

Sequential Data!

Imagine Your Boss Wants You to Solve a Prolem Like...

Sequence of words

Automatically generate caption with the given image

Predict whether a company would be bankrupted

Translate one sentence into another language

Classify whether the word is owns` name or not

Sequential Data!

Imagine Your Boss Wants You to Solve a Problem Like...

Automatically generate caption with the given image

Sequence of balance

Predict whether a company would be bankrupted

Translate one sentence into another language

Classify whether the word is owns` name or not

Sequential Data!

Imagine Your Boss Wants You to Solve a Problem Like...

Automatically generate caption with the given image

Predict whether a company would be bankrupted

Sequence of words

Translate one sentence into another language

Classify whether the word is owns` name or not

Sequential Data!

Imagine Your Boss Wants You to Solve a Problem Like...

Automatically generate caption with the given image

Predict whether a company would be bankrupted

Translate one sentence into another language

Sequence of words

Classify whether the word is owns` name or not

Sequential Data!

Imagine Your Boss Wants You to Solve a Problem Like...

Automatically generate caption with the given image

Predict whether a company would be bankrupted

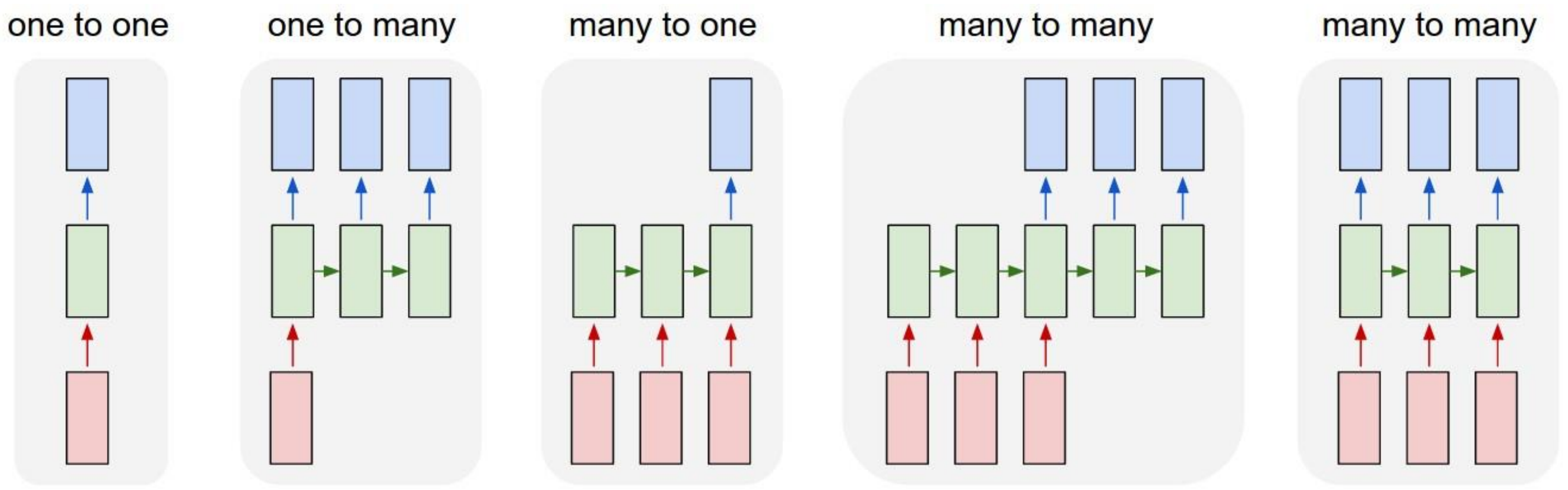
Translate one sentence into another language

Sequence of words

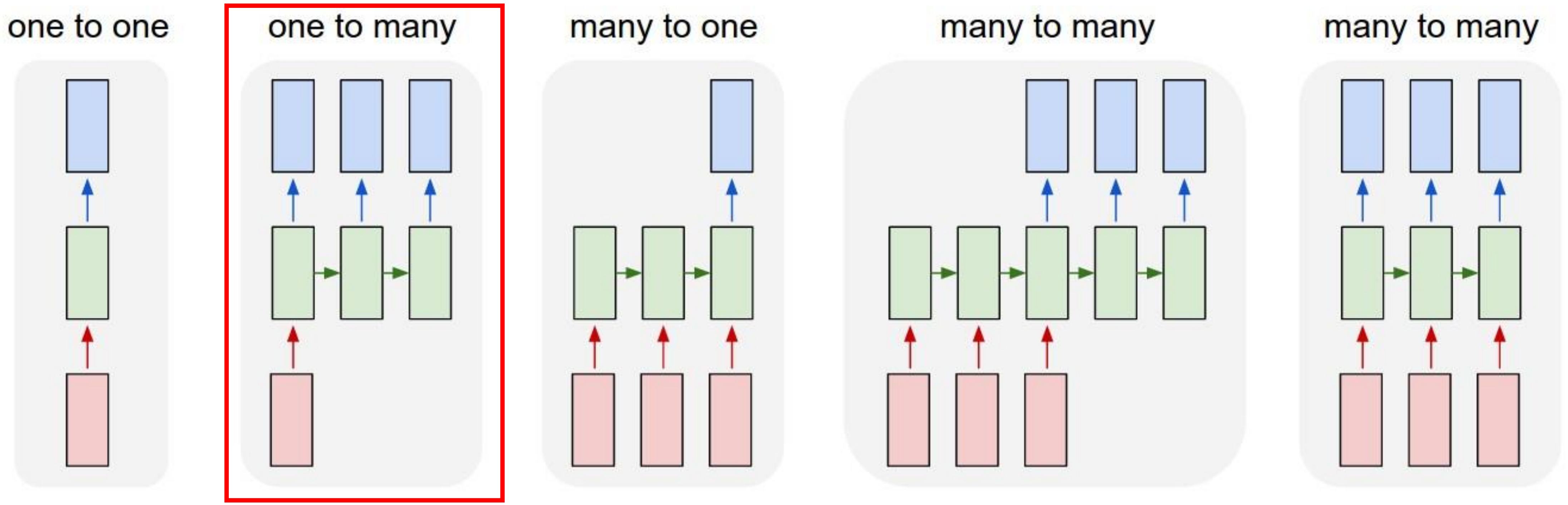
Classify whether the word is owns` name or not

Sequential Data!

Types of Task Dealing with Sequential Data

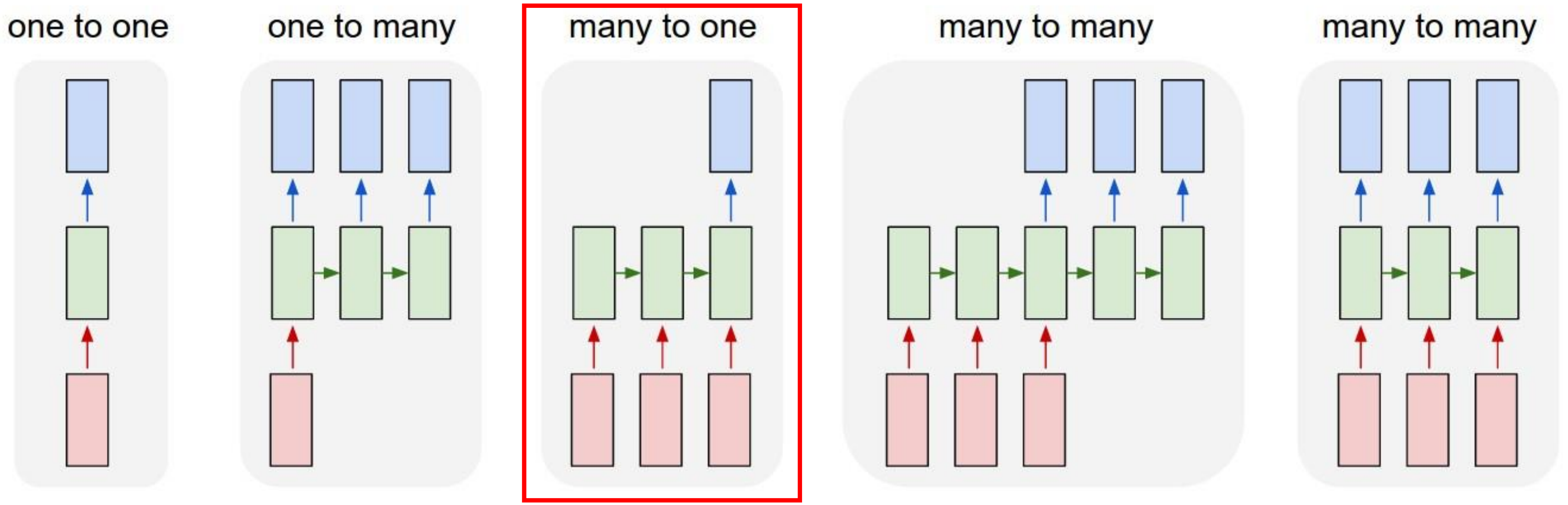


Types of Task Dealing with Sequential Data



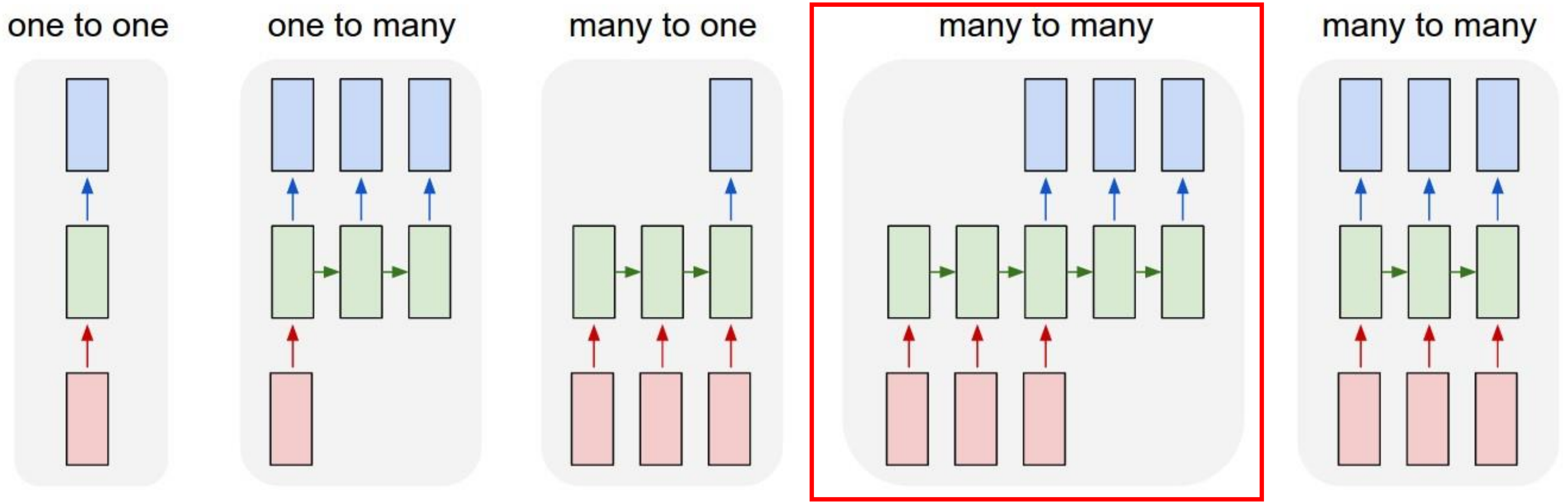
Automatically generate caption with the given image

Types of Task Dealing with Sequential Data



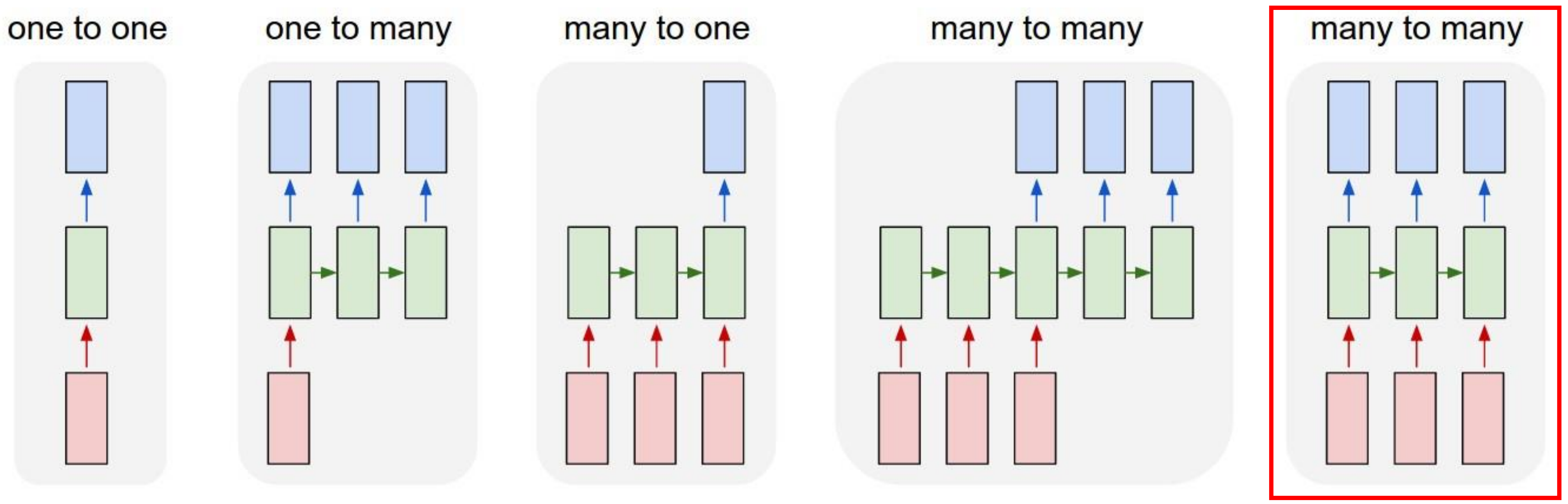
Predict whether a company would be bankrupted

Types of Task Dealing with Sequential Data



Translate one sentence into another language

Types of Task Dealing with Sequential Data



Classify whether the word is owns` name or not

http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Classical Approach for Time Series Analysis

Classical Approach for Time Series Analysis

- Time domain analysis
- Frequency domain analysis
- Nearest neighbors analysis
- Probabilistic Model
- (S)AR(I)MA(X) models
- Decomposition
- Nonlinear Dynamics
- Machine Learning

https://medium.com/@alexrachnog/deep-learning-the-final-frontier-for-signal-processing-and-time-series-analysis-734307167ad6

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis
- Nearest neighbors analysis
- Probabilistic Model
- (S)AR(I)MA(X) models
- Decomposition
- Nonlinear Dynamics
- Machine Learning

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis
- Probabilistic Model
- (S)AR(I)MA(X) models
- Decomposition
- Nonlinear Dynamics
- Machine Learning

https://medium.com/@alexrachnog/deep-learning-the-final-frontier-for-signal-processing-and-time-series-analysis-734307167ad6

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis $\longrightarrow$ Dynamic time warping (DTW)
- Probabilistic Model
- (S)AR(I)MA(X) models
- Decomposition
- Nonlinear Dynamics
- Machine Learning

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis $\longrightarrow$ Dynamic time warping (DTW)
- Probabilistic Model $\longrightarrow$ Language modeling
- (S)AR(I)MA(X) models
- Decomposition
- Nonlinear Dynamics
- Machine Learning

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis $\longrightarrow$ Dynamic time warping (DTW)
- Probabilistic Model $\longrightarrow$ Language modeling
- (S)AR(I)MA(X) models $\longrightarrow$ Autocorrelation inside of time series
- Decomposition
- Nonlinear Dynamics
- Machine Learning

https://medium.com/@alexrachnog/deep-learning-the-final-frontier-for-signal-processing-and-time-series-analysis-734307167ad6

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis $\longrightarrow$ Dynamic time warping (DTW)
- Probabilistic Model $\longrightarrow$ Language modeling
- (S)AR(I)MA(X) models $\longrightarrow$ Autocorrelation inside of time series
- Decomposition $\longrightarrow$ Time series = trend part + seasonal part + residuals
- Nonlinear Dynamics
- Machine Learning

https://medium.com/@alexrachnog/deep-learning-the-final-frontier-for-signal-processing-and-time-series-analysis-734307167ad6

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis $\longrightarrow$ Dynamic time warping (DTW)
- Probabilistic Model $\longrightarrow$ Language modeling
- (S)AR(I)MA(X) models $\longrightarrow$ Autocorrelation inside of time series
- Decomposition $\longrightarrow$ Time series = trend part + seasonal part + residuals
- Nonlinear Dynamics $\longrightarrow$ Differential Equation (ordinary, partial, stochastic, etc..)
- Machine Learning

Classical Approach for Time Series Analysis

- Time domain analysis $\longrightarrow$ width, step, height of signal
- Frequency domain analysis $\longrightarrow$ Fourier analysis or wavelets
- Nearest neighbors analysis $\longrightarrow$ Dynamic time warping (DTW)
- Probabilistic Model $\longrightarrow$ Language modeling
- (S)AR(I)MA(X) models $\longrightarrow$ Autocorrelation inside of time series
- Decomposition $\longrightarrow$ Time series = trend part + seasonal part + residuals
- Nonlinear Dynamics $\longrightarrow$ Differential Equation (ordinary, partial, stochastic, etc..)
- Machine Learning $\longrightarrow$ Use ML model with hand-made features

Deep Learning Dealing with Sequential Data

Deep Learning Dealing with Sequential Data

MLP? Stack of fully connected layers

CNN? Stack of (convolution + pooling + fully connected) layers

Deep Learning Dealing with Sequential Data

MLP? Stack of fully connected layers

Cannot handle a sequence with arbitrary length

For fixed length sequence, require lots of parameters

CNN? Stack of (convolution + pooling + fully connected) layers

Deep Learning Dealing with Sequential Data

MLP? Stack of fully connected layers

Cannot handle a sequence with arbitrary length

For fixed length sequence, require lots of parameters

CNN? Stack of (convolution + pooling + fully connected) layers

Actually perform quite well on time series analysis

Recommend to read: https://machinelearningmastery.com/how-to-develop-convolutional-neural-networks-for-multi-step-time-series-forecasting/

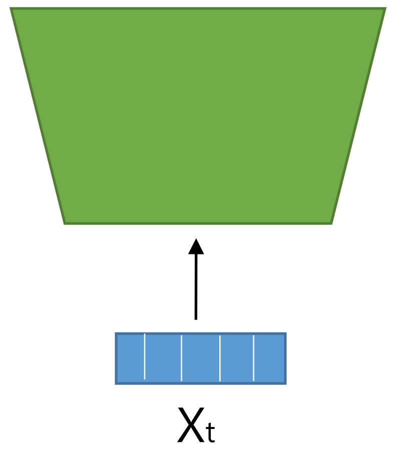
Recurrent Neural Network

Recurrent Neural Network

Process both new inputs and model output of previous input!

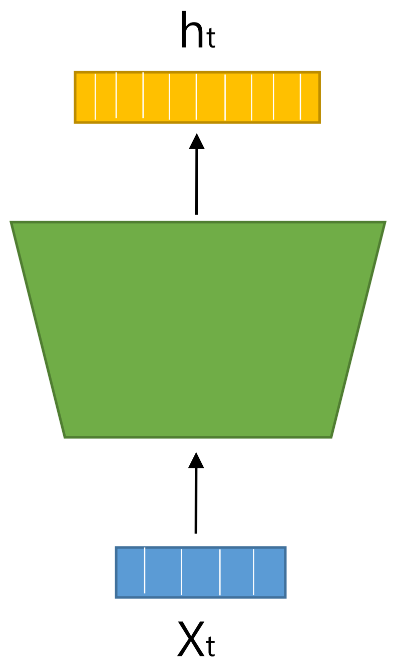
Recurrent Neural Network

Process both new inputs and model output of previous input!



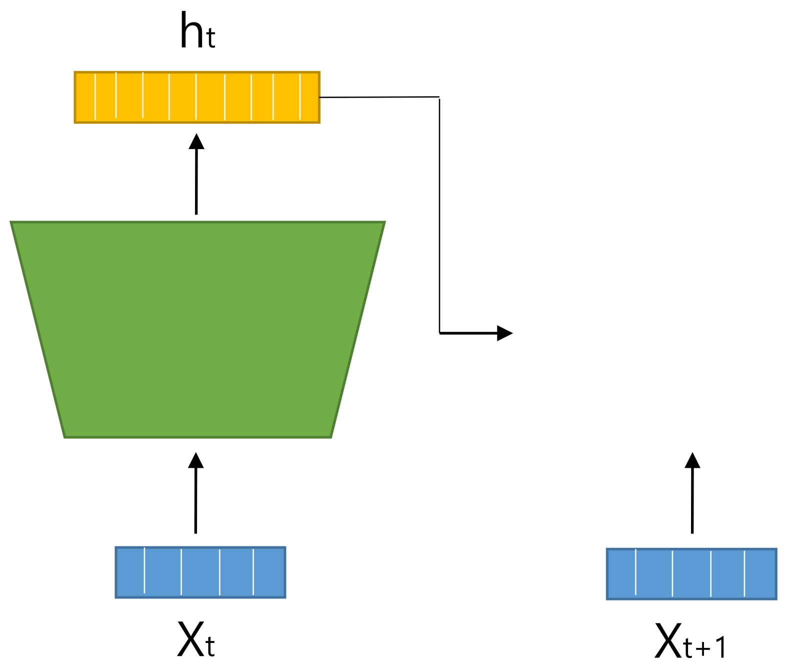
Recurrent Neural Network

Process both new inputs and model output of previous input!



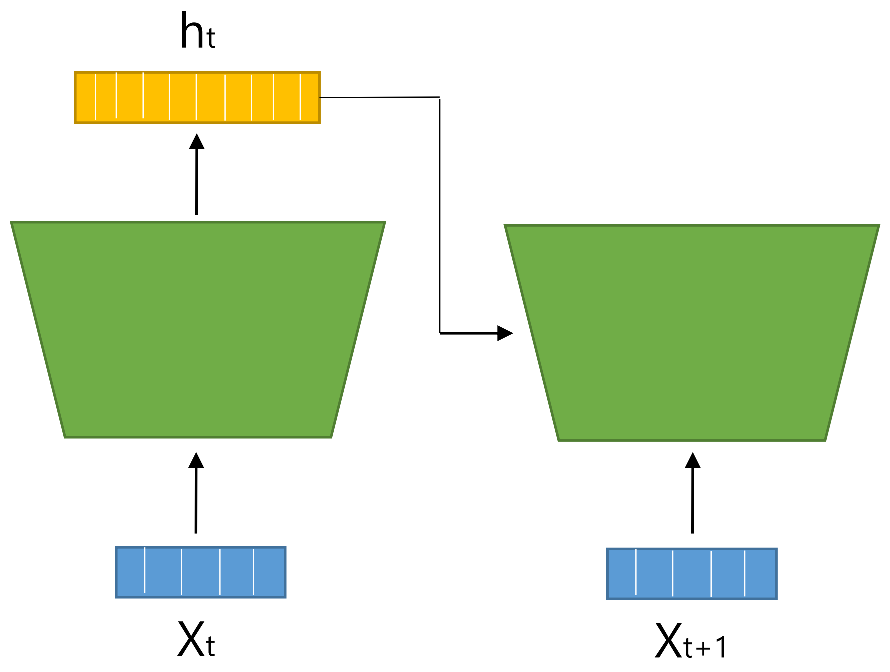
Recurrent Neural Network

Process both new inputs and model output of previous input!



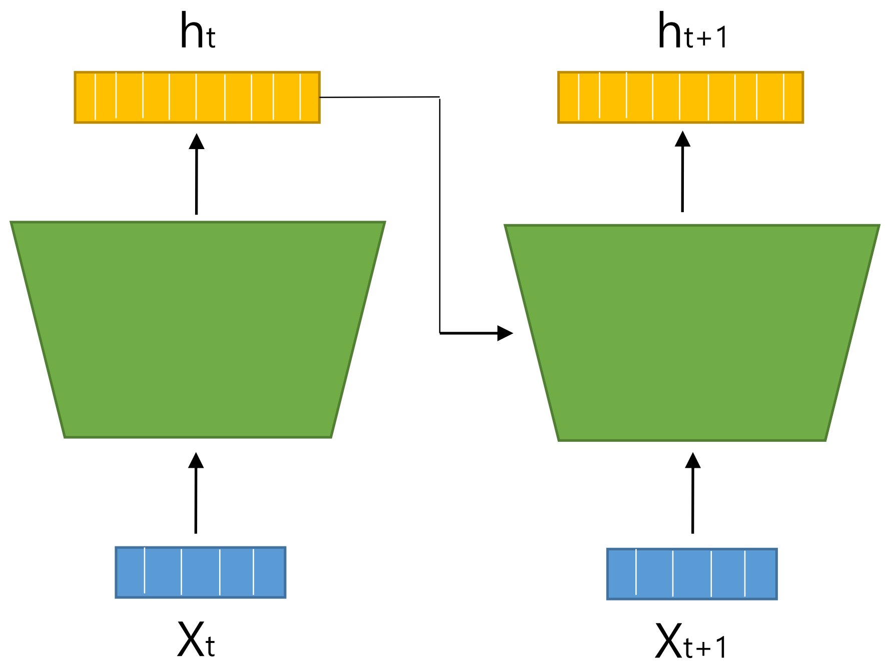
Recurrent Neural Network

Process both new inputs and model output of previous input!



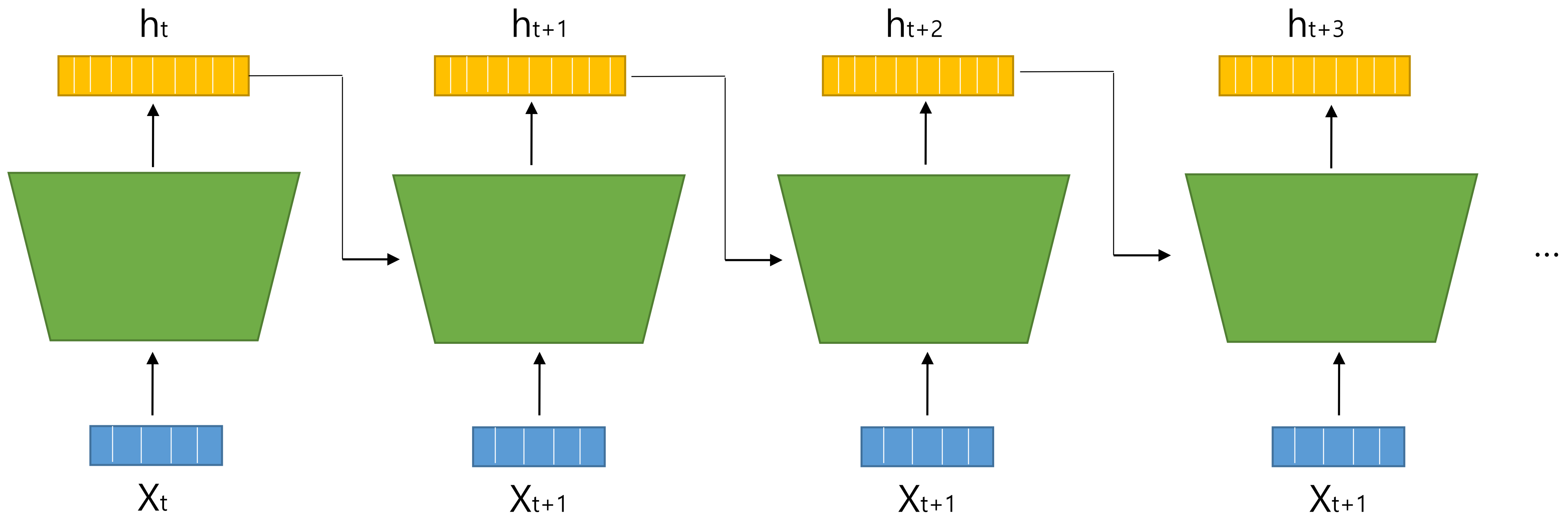
Recurrent Neural Network

Process both new inputs and model output of previous input!



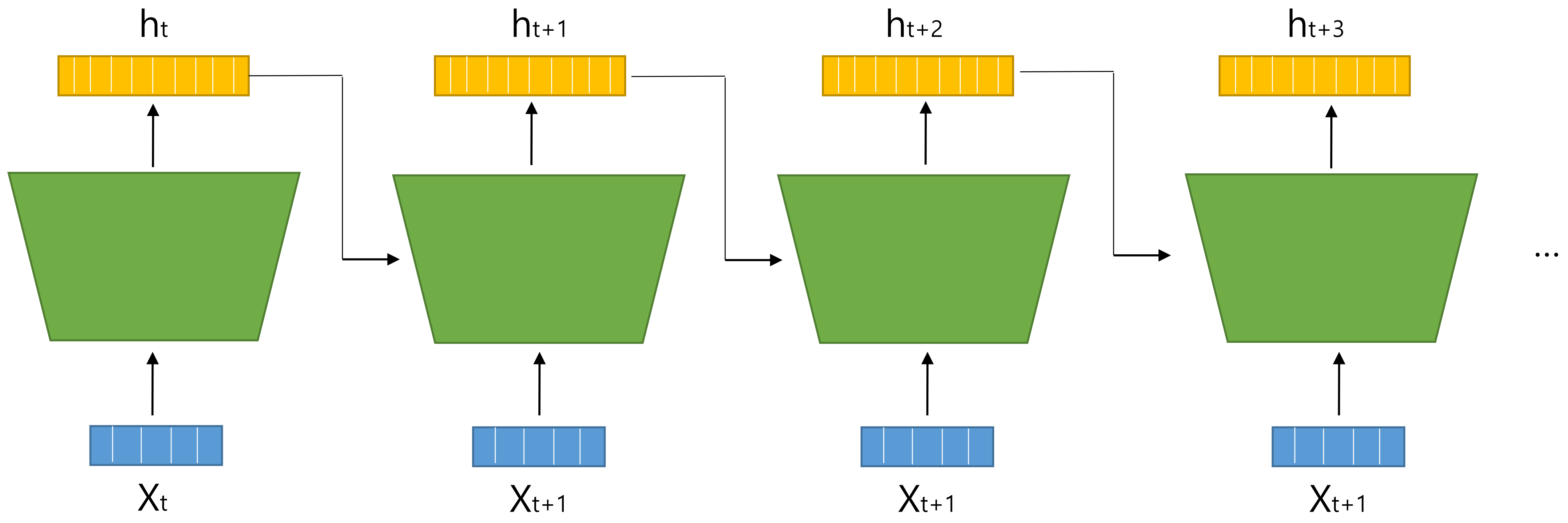
Recurrent Neural Network

Process both new inputs and model output of previous input!



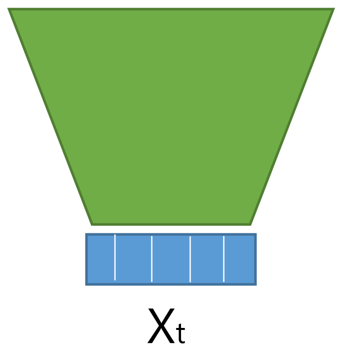
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



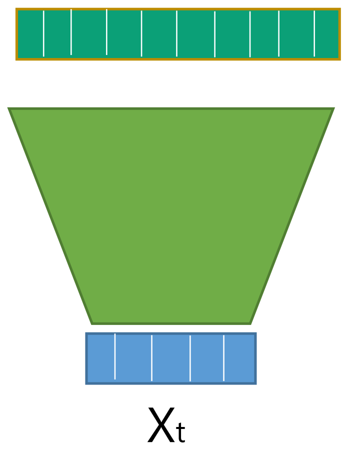
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



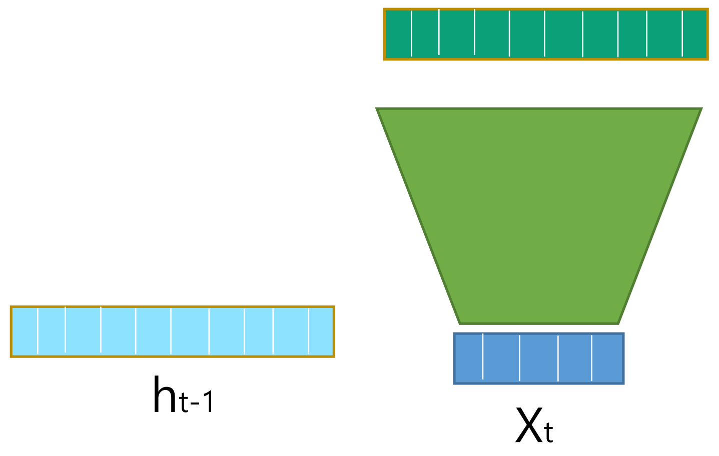
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



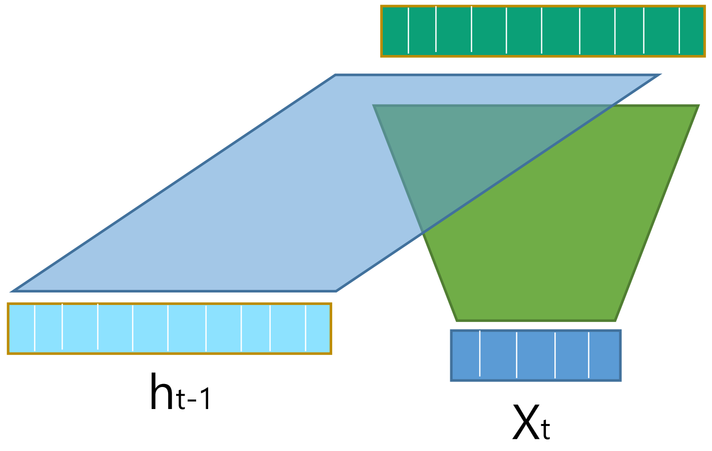
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



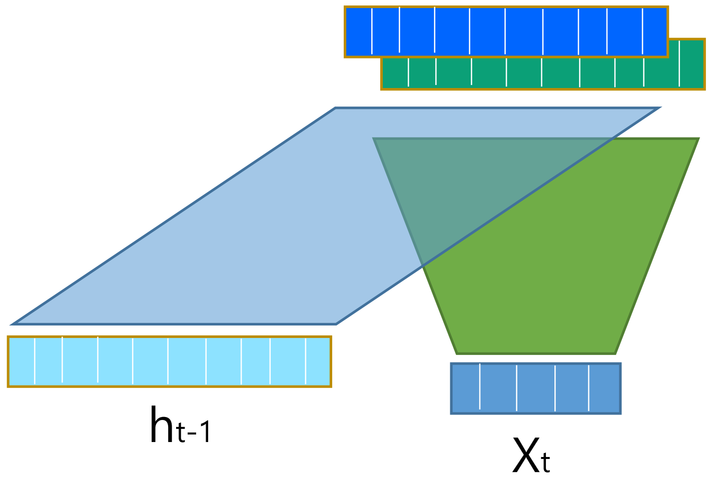
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



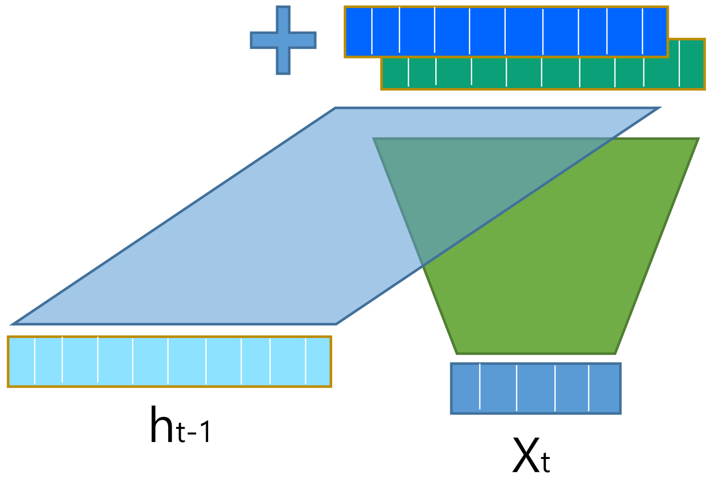
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



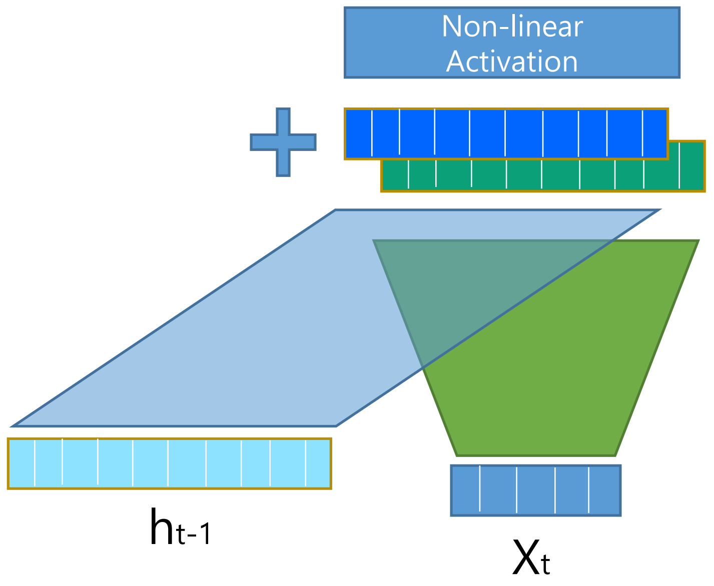
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



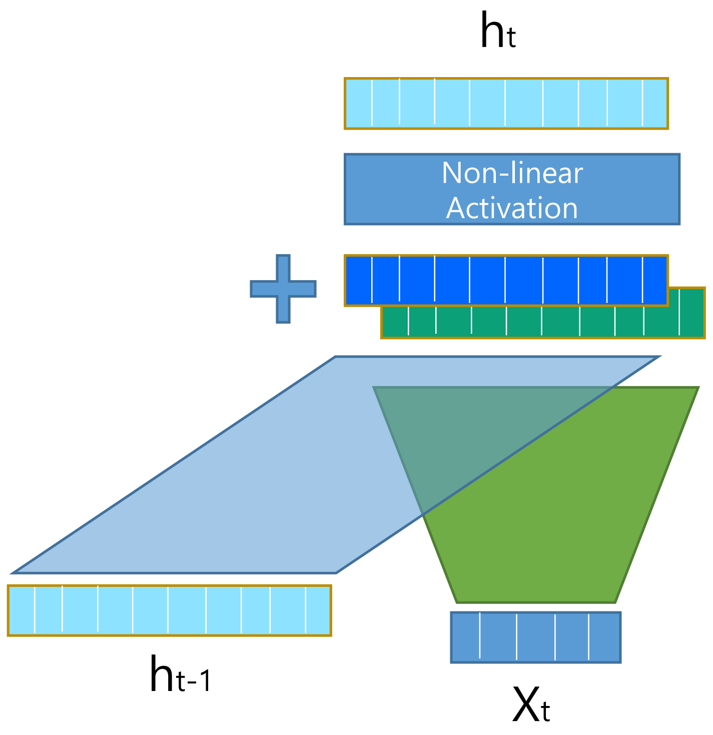
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



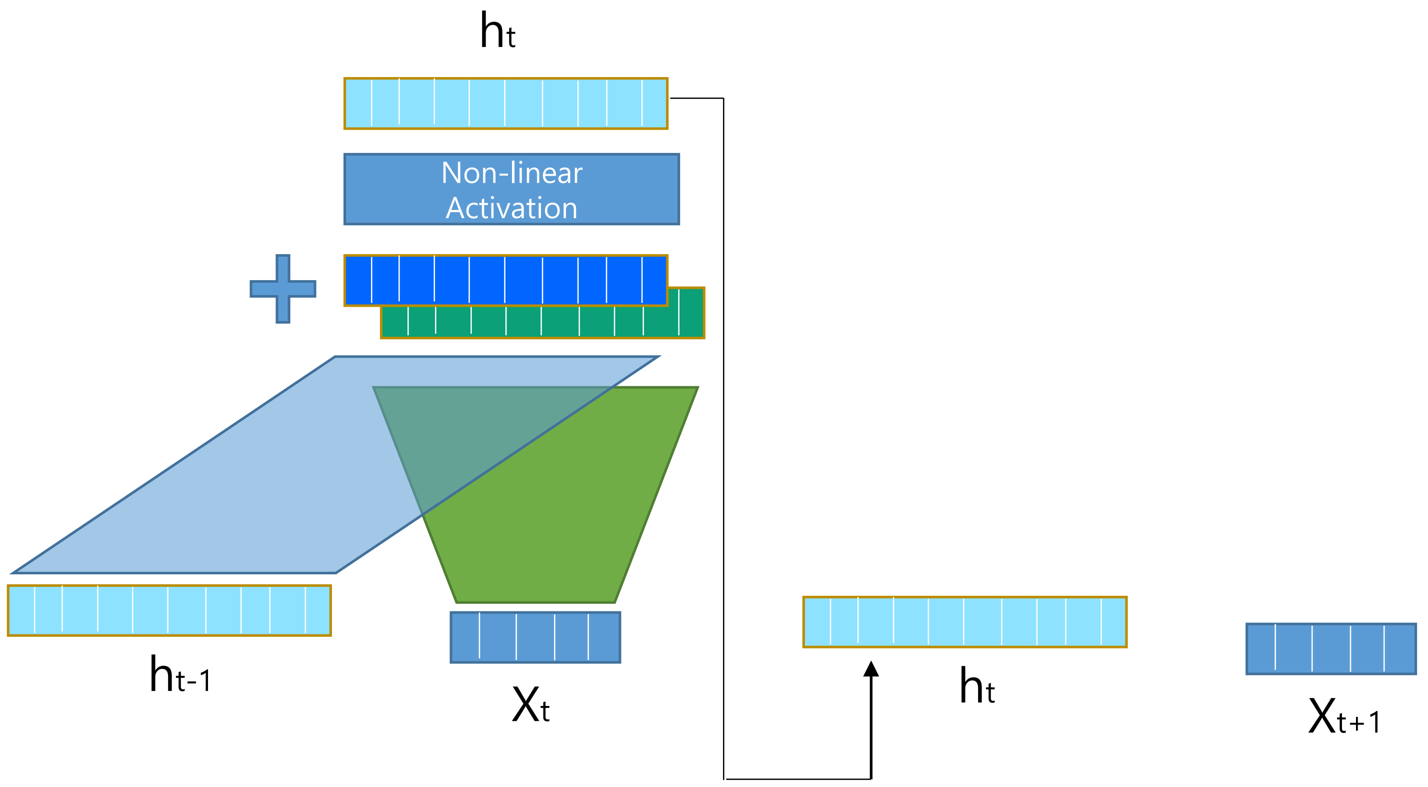
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



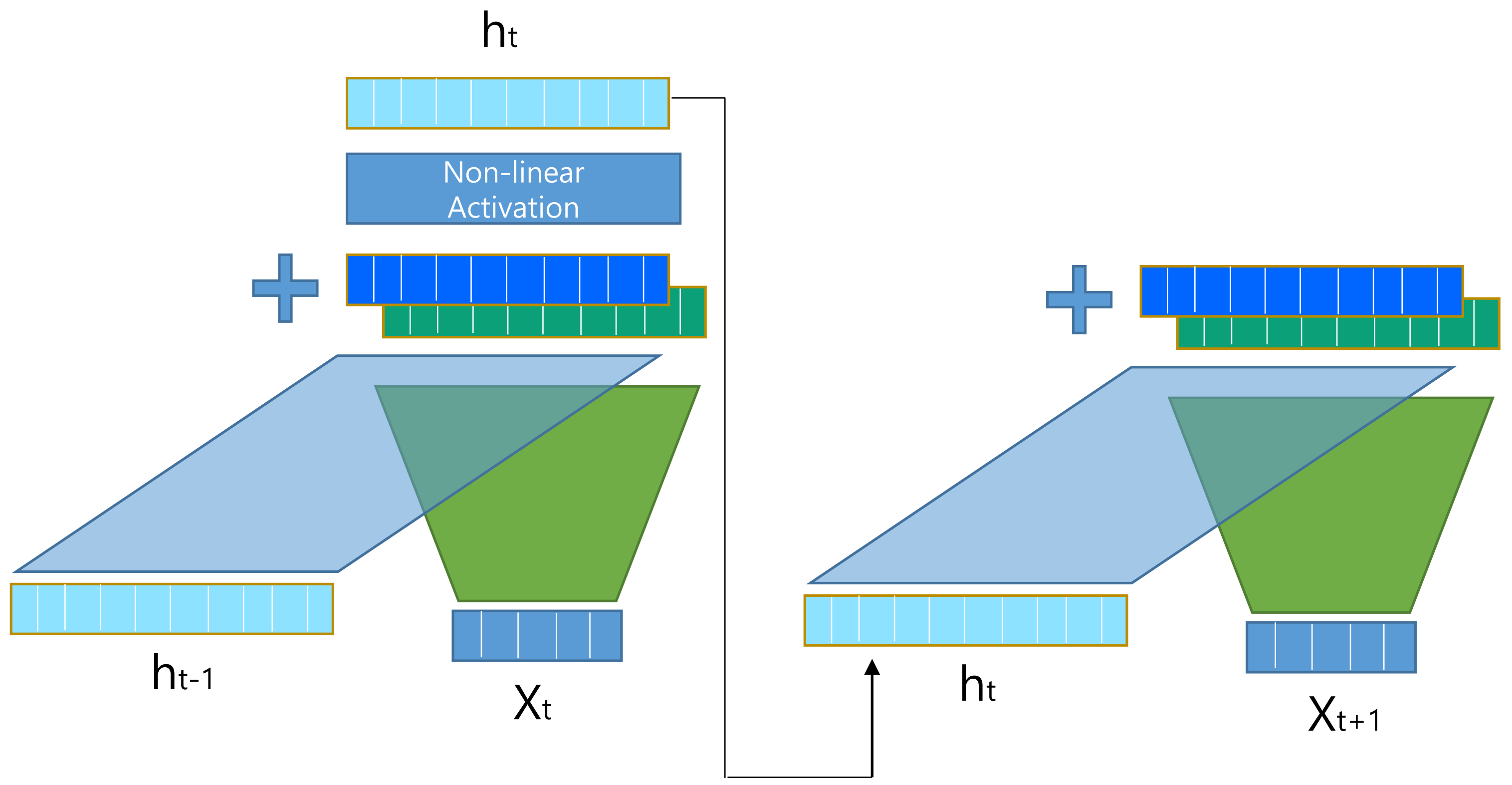
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



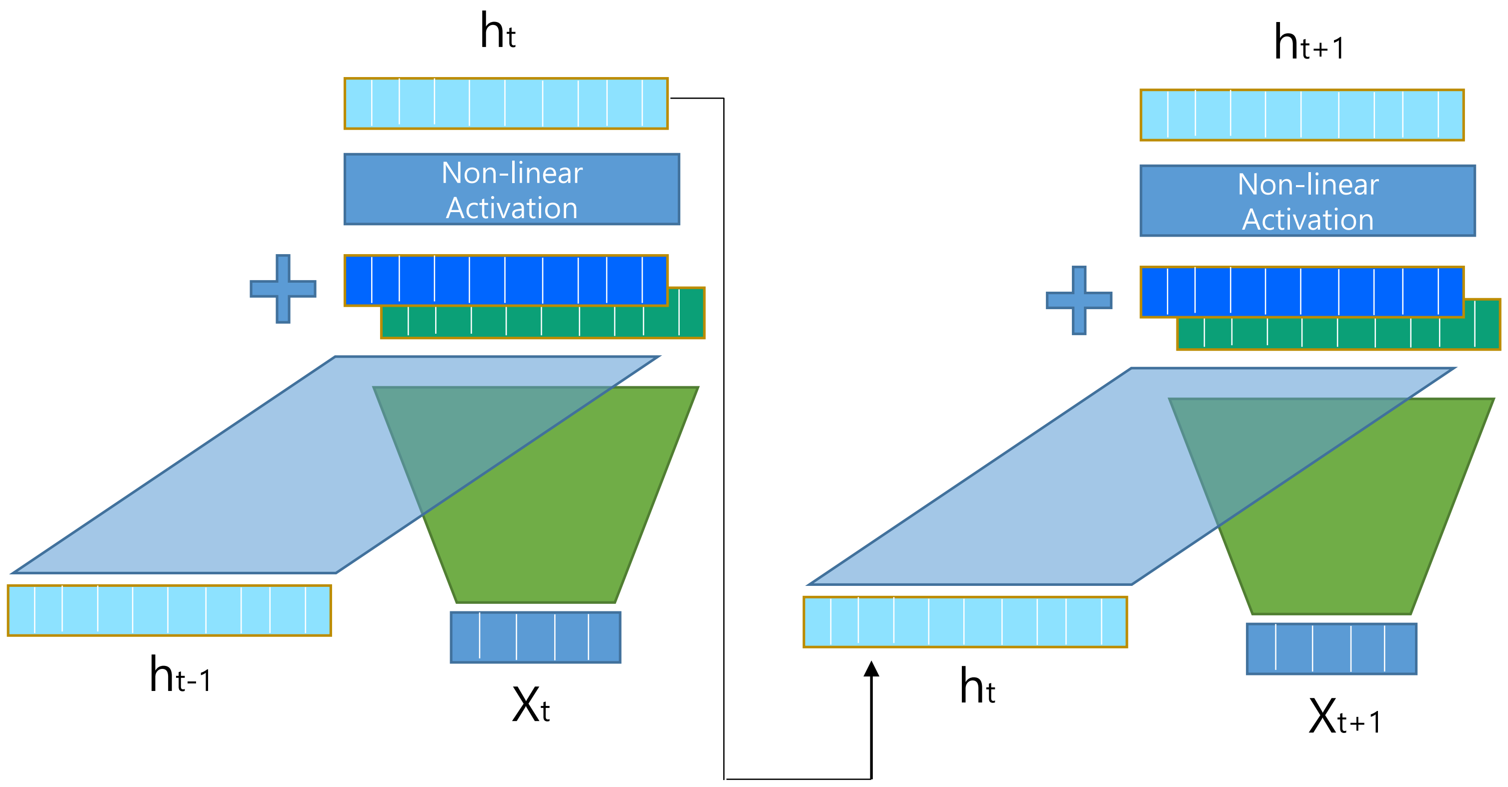
Recurrent Neural Network

But exactly, how can we combine new input and previous output?



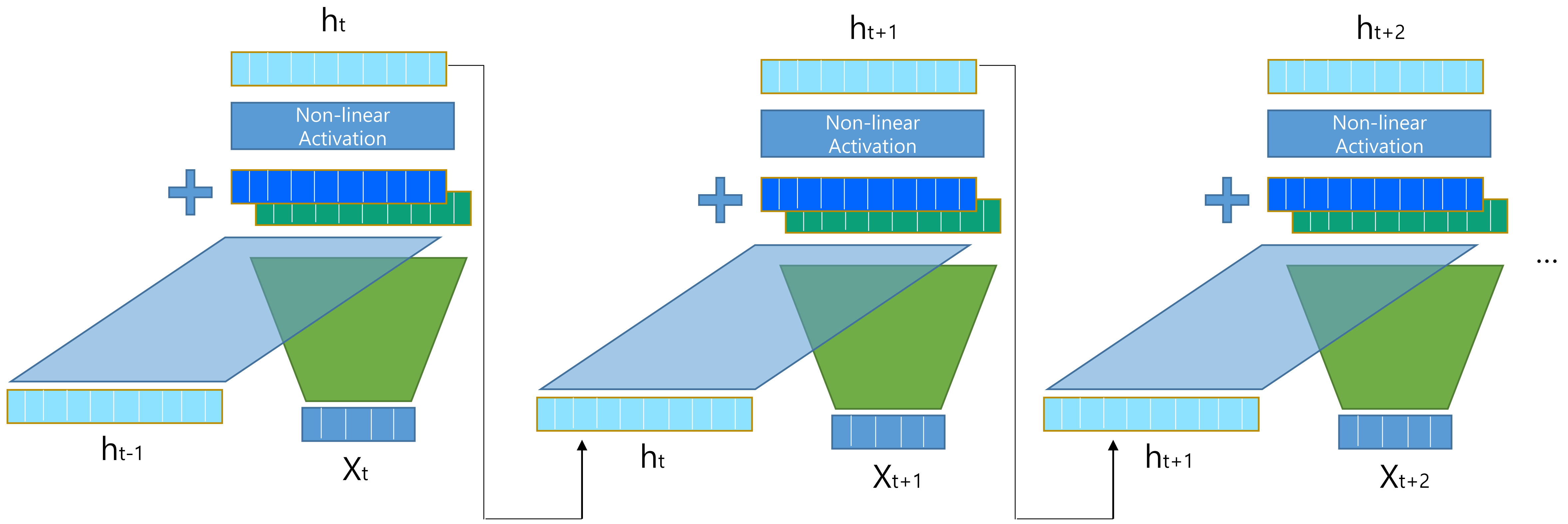
Recurrent Neural Network

But exactly, how can we combine new input and previous output?

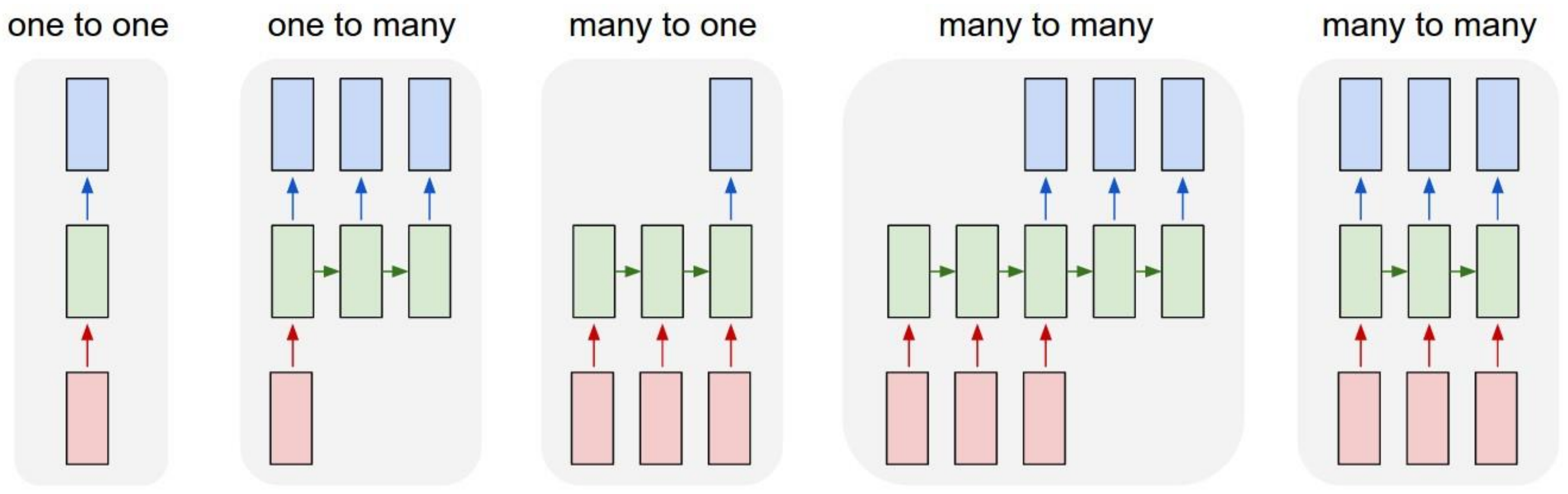


Recurrent Neural Network

But exactly, how can we combine new input and previous output?

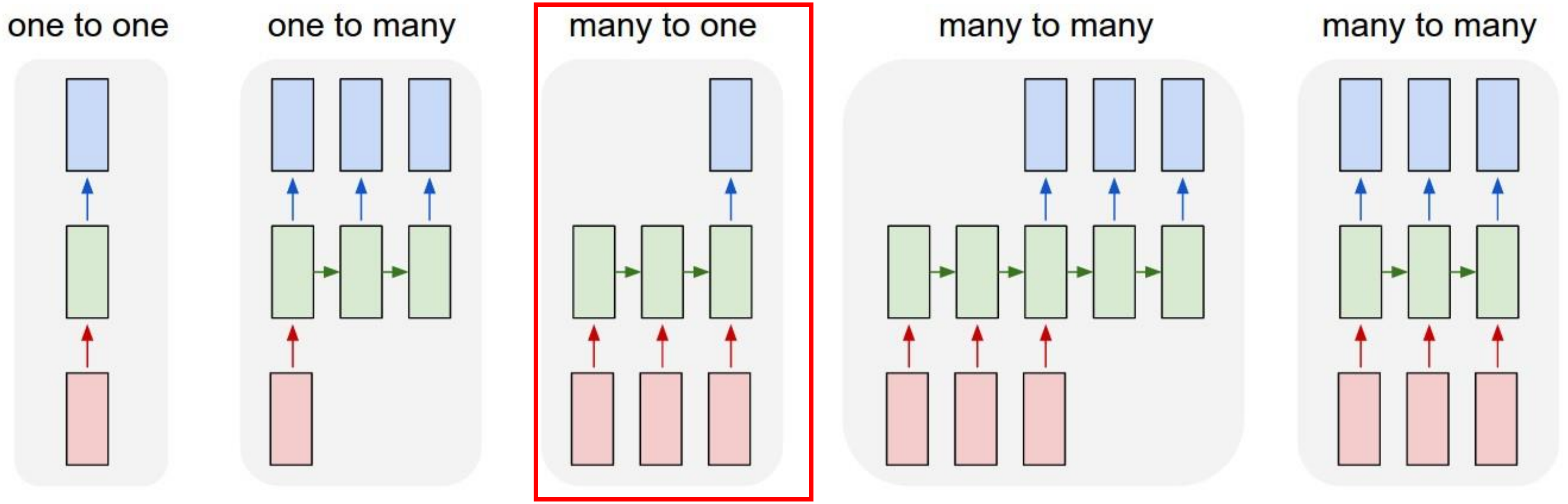


Types of Task Dealing with Sequential Data



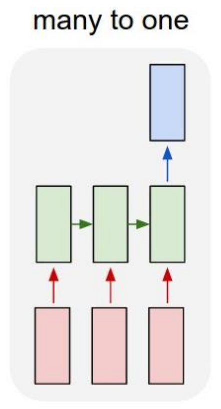
http://karpathy.github.io/2015/05/21/rnn-effectiveness/

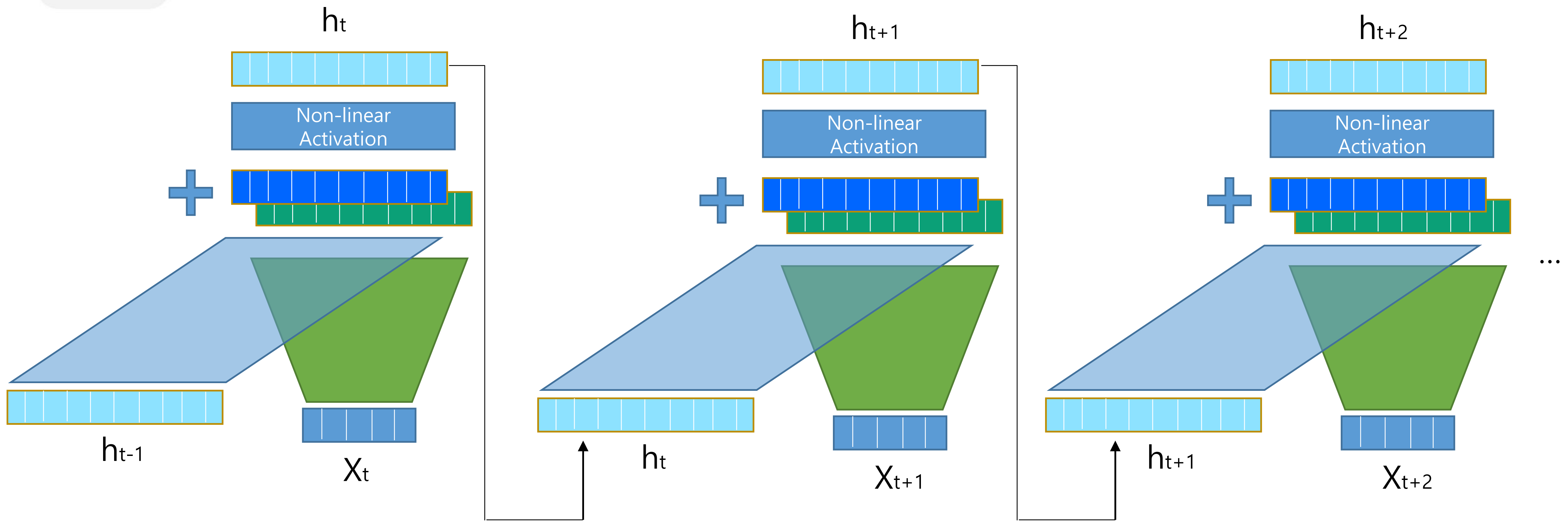
Types of Task Dealing with Sequential Data



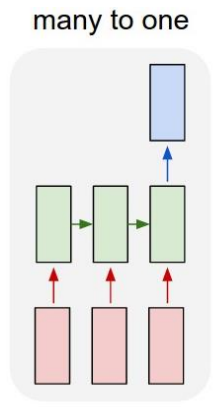
http://karpathy.github.io/2015/05/21/rnn-effectiveness/

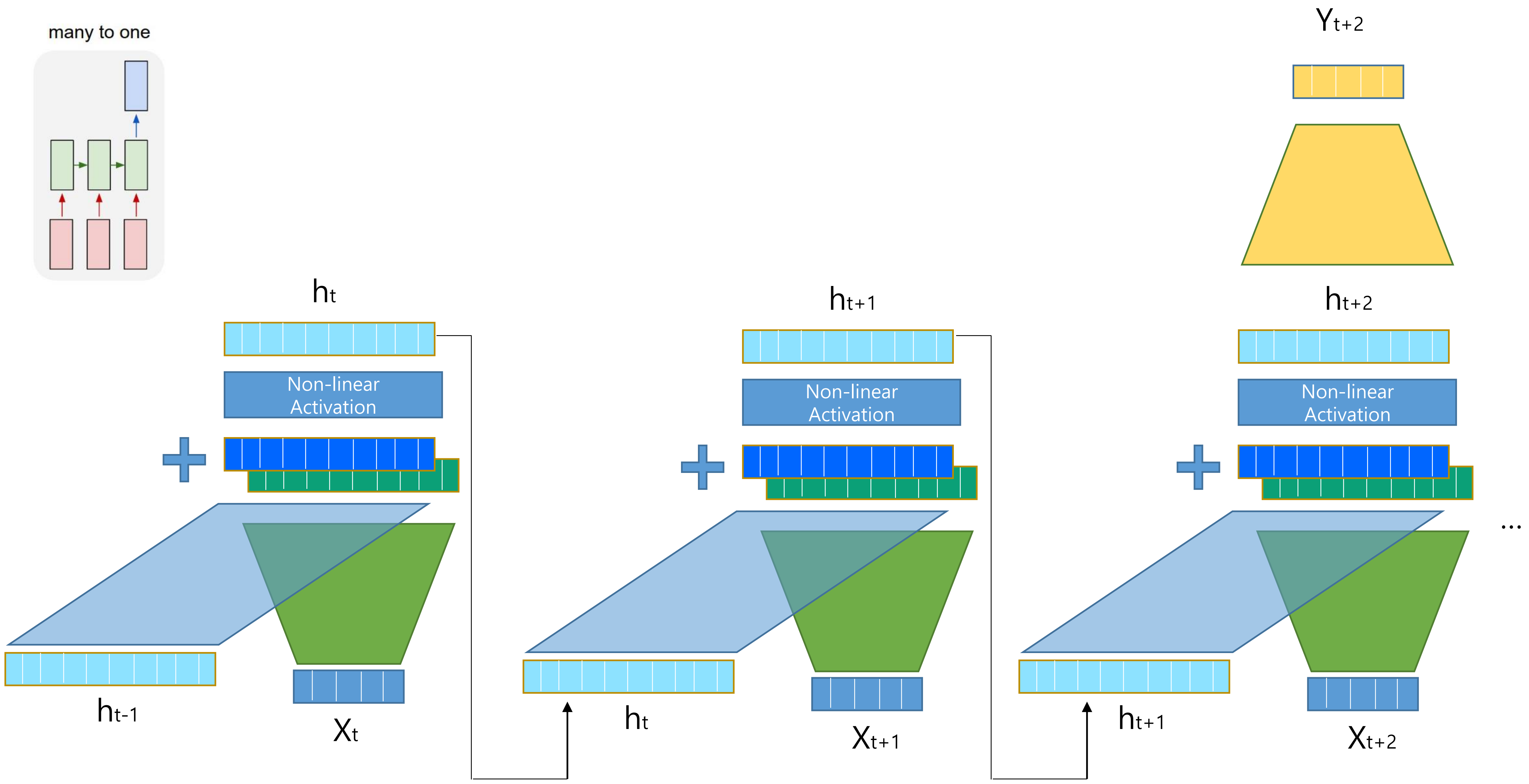
Recurrent Neural Network



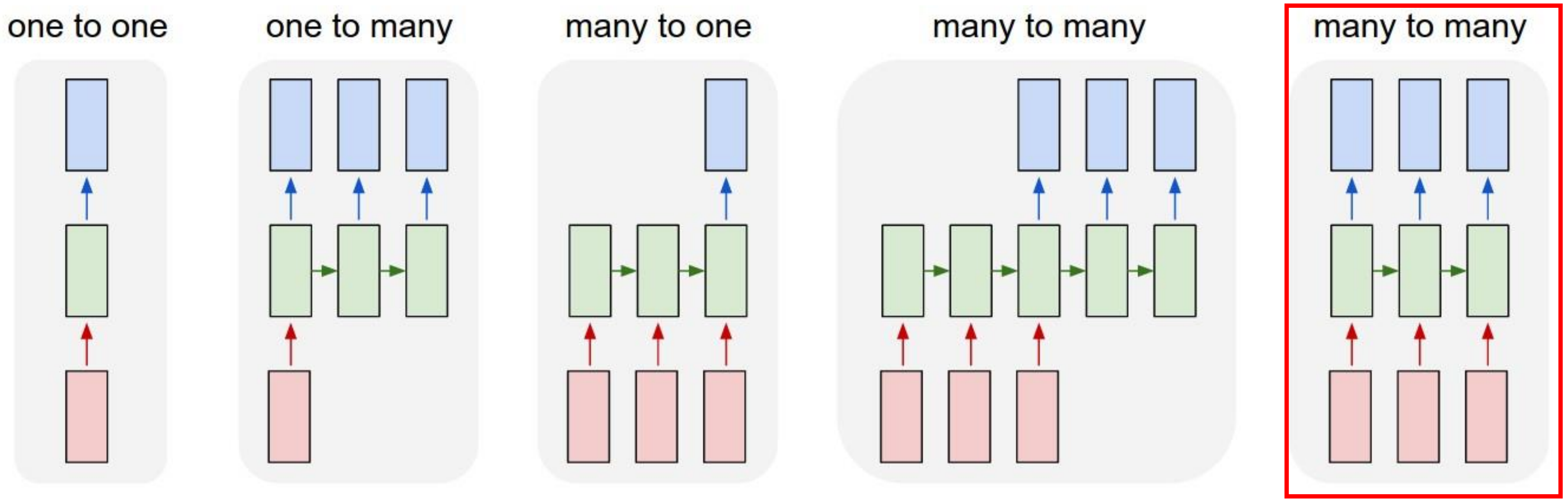


Recurrent Neural Network

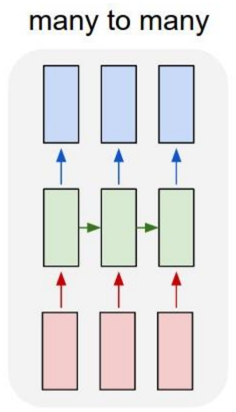


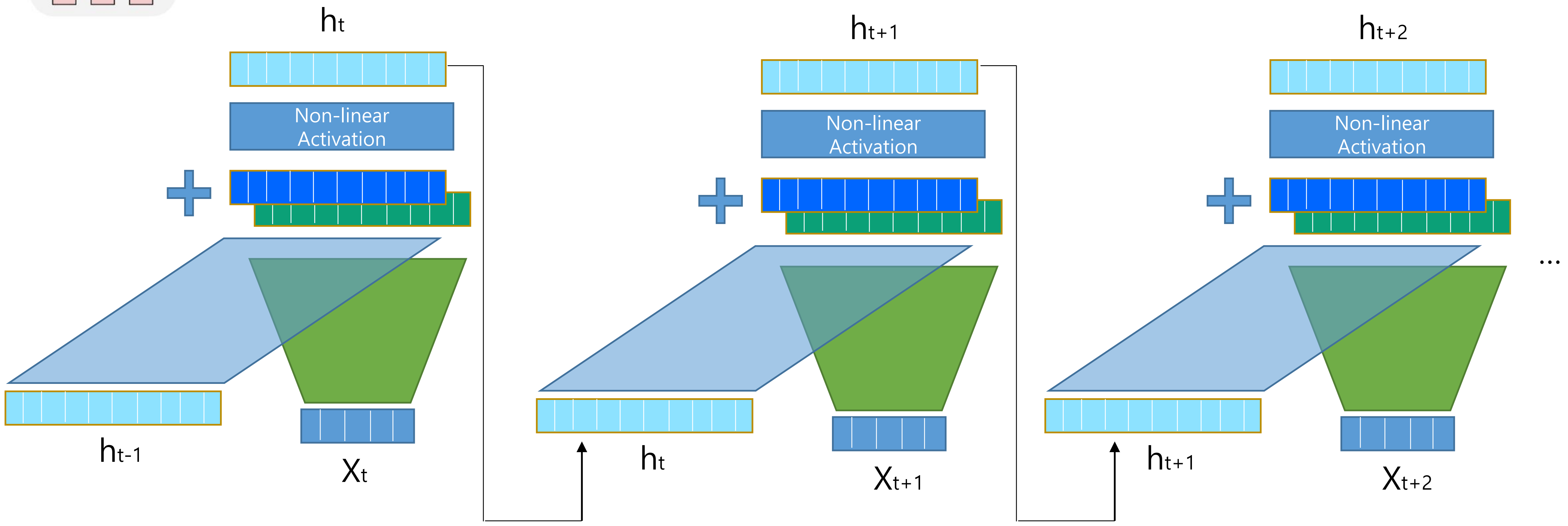


Types of Task Dealing with Sequential Data

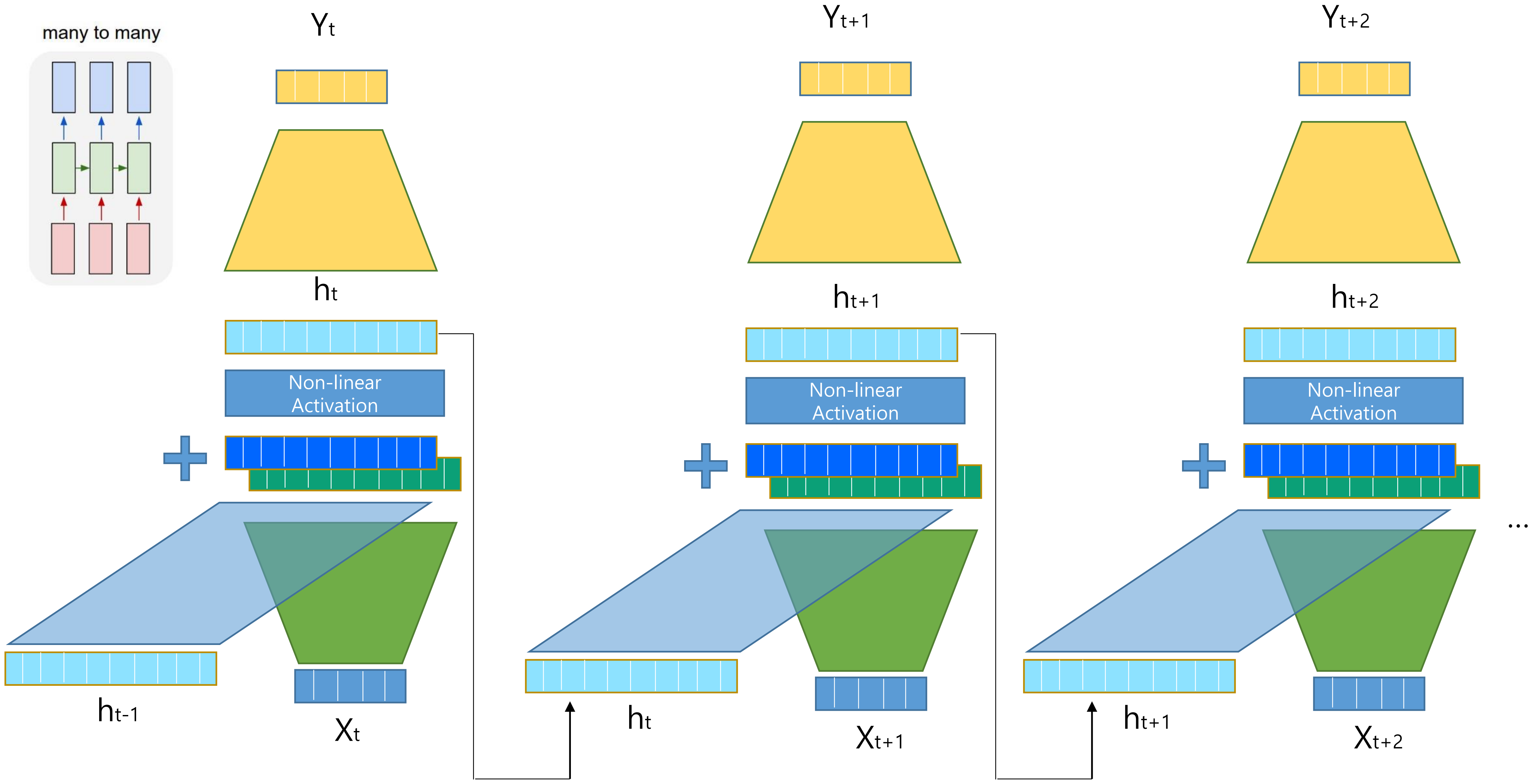


Recurrent Neural Network

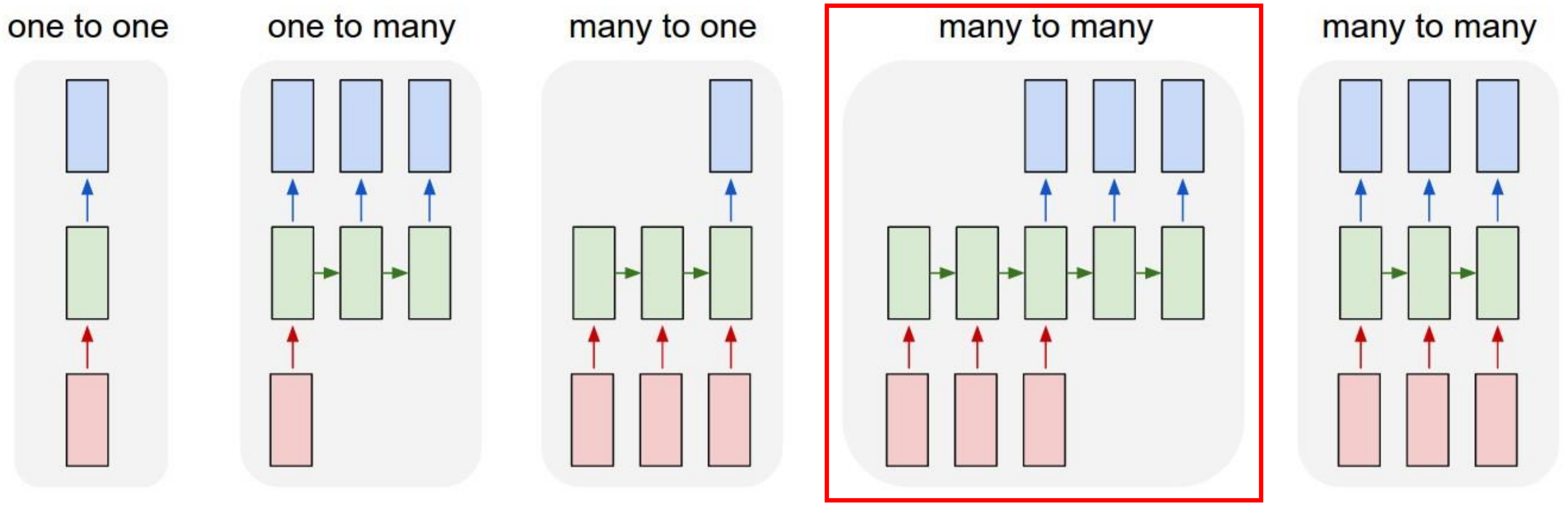




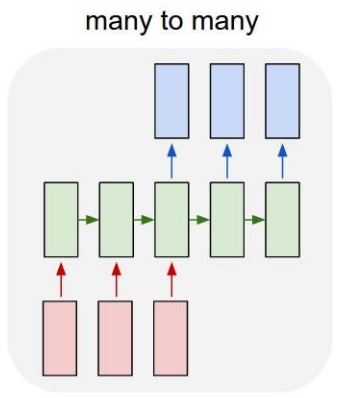
Recurrent Neural Network

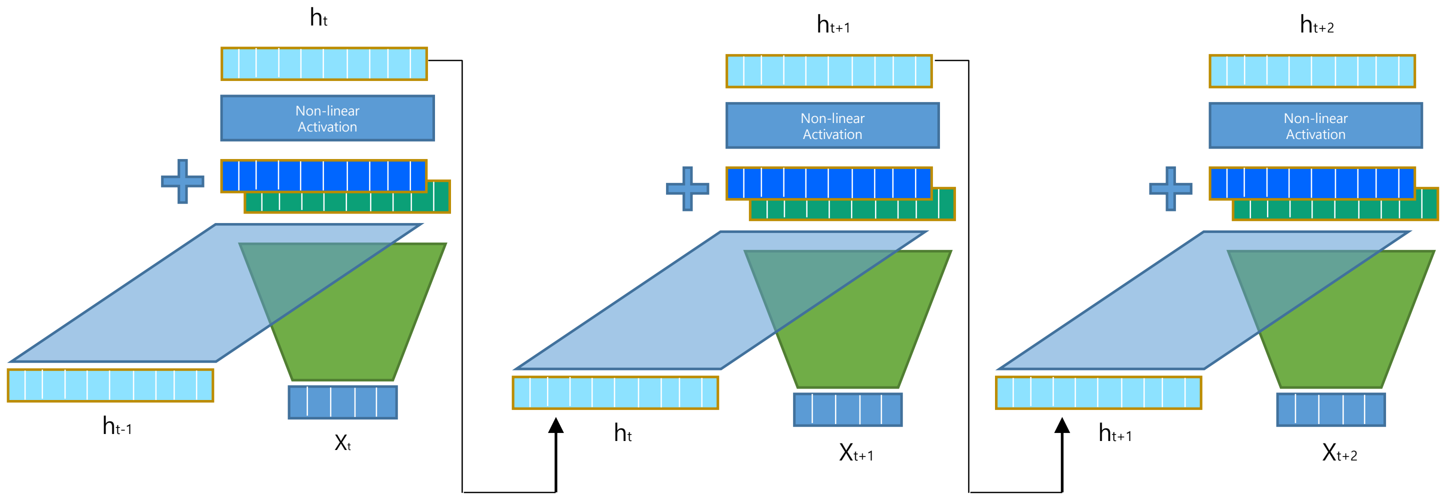


Types of Task Dealing with Sequential Data

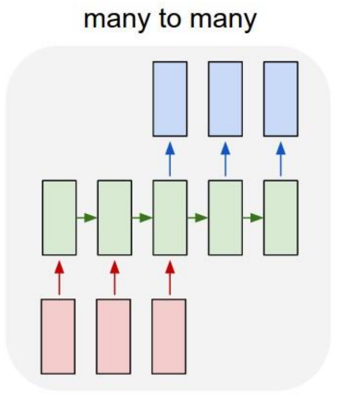


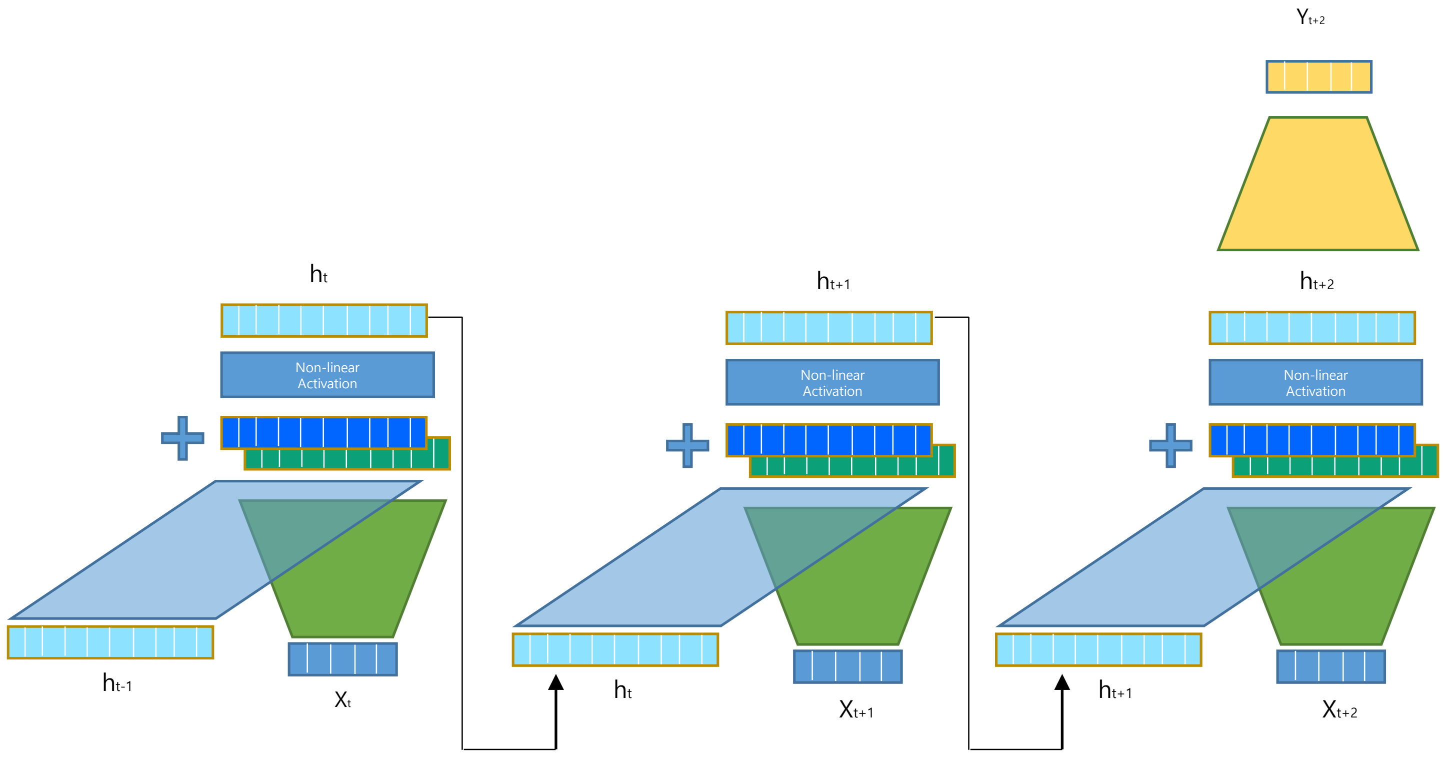
Recurrent Neural Network



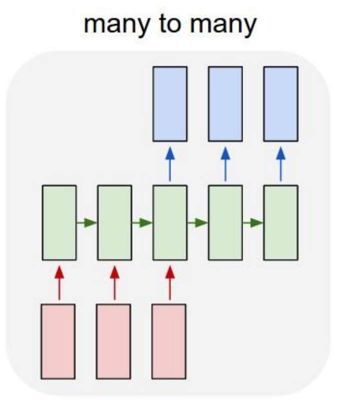


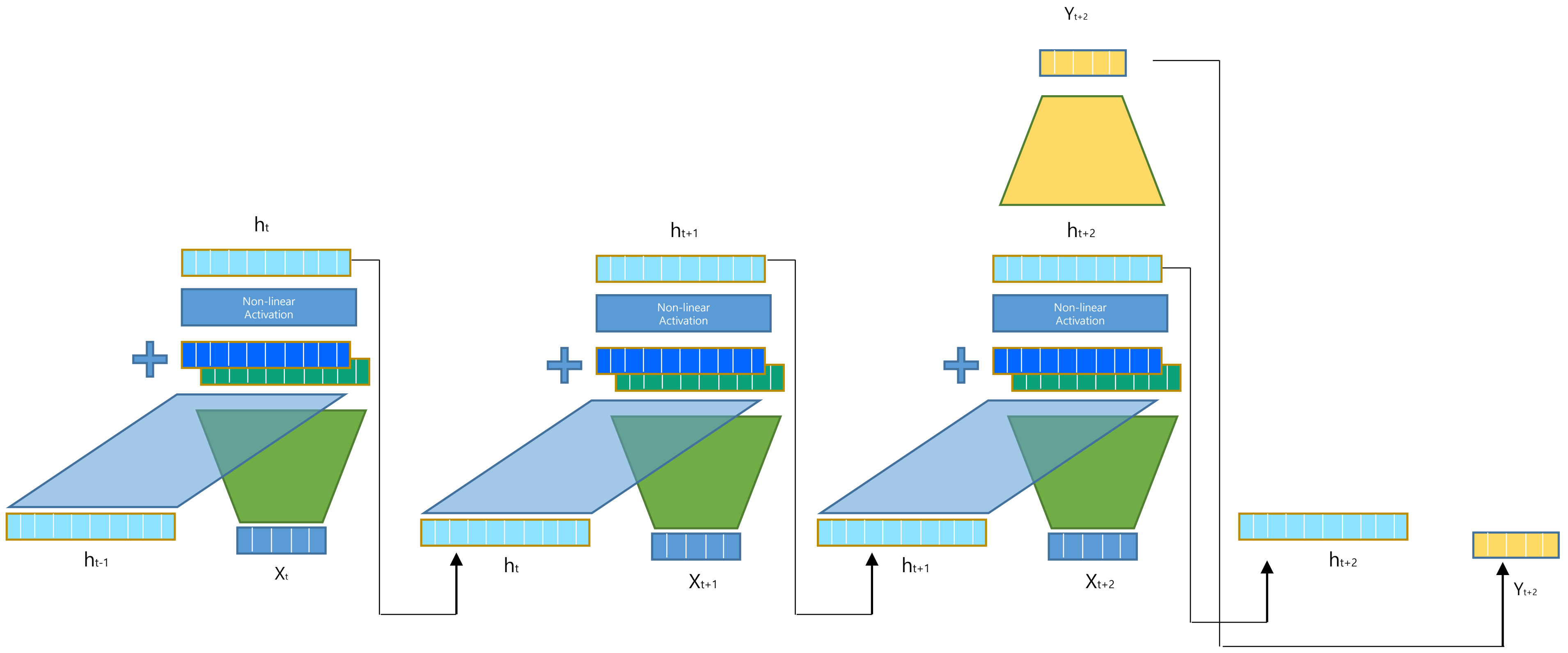
Recurrent Neural Network



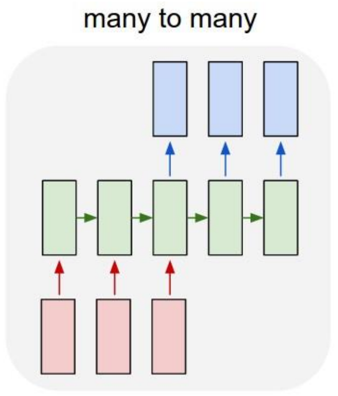


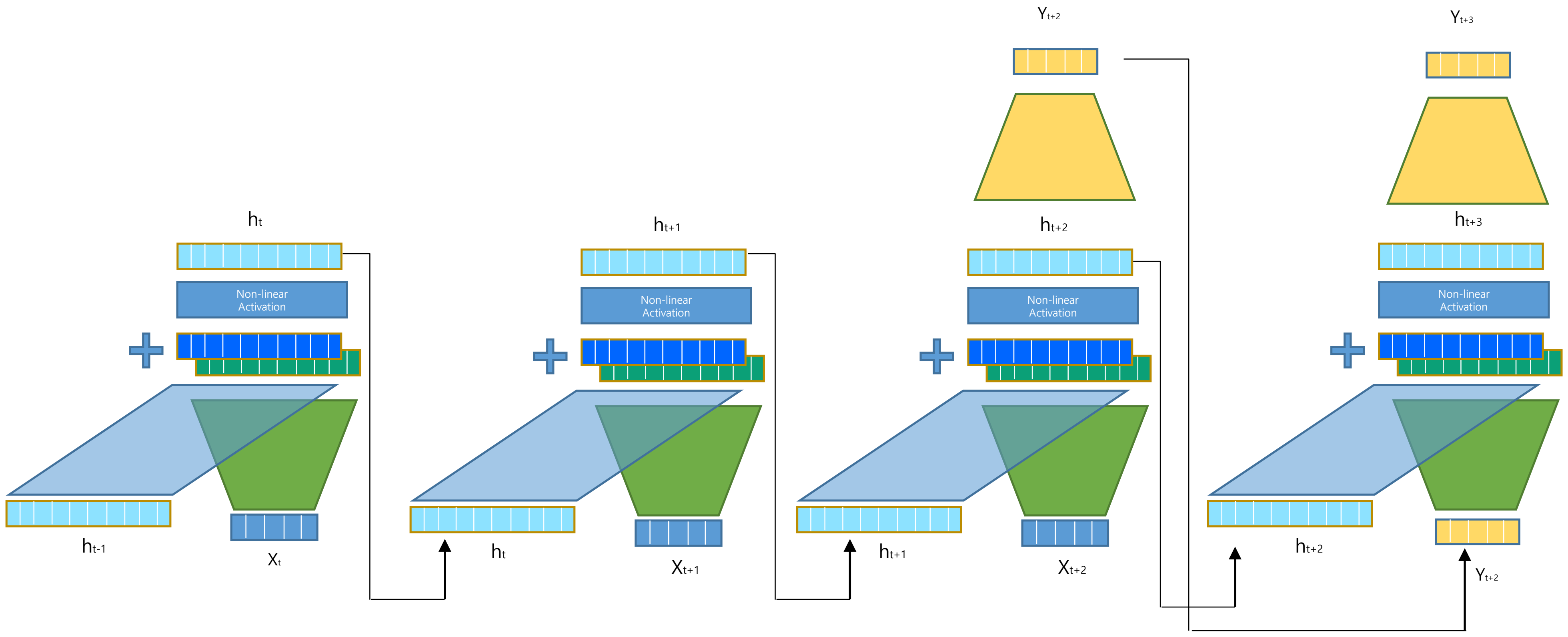
Recurrent Neural Network





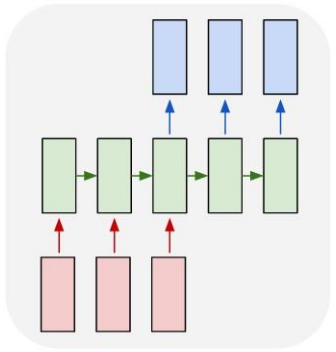
Recurrent Neural Network

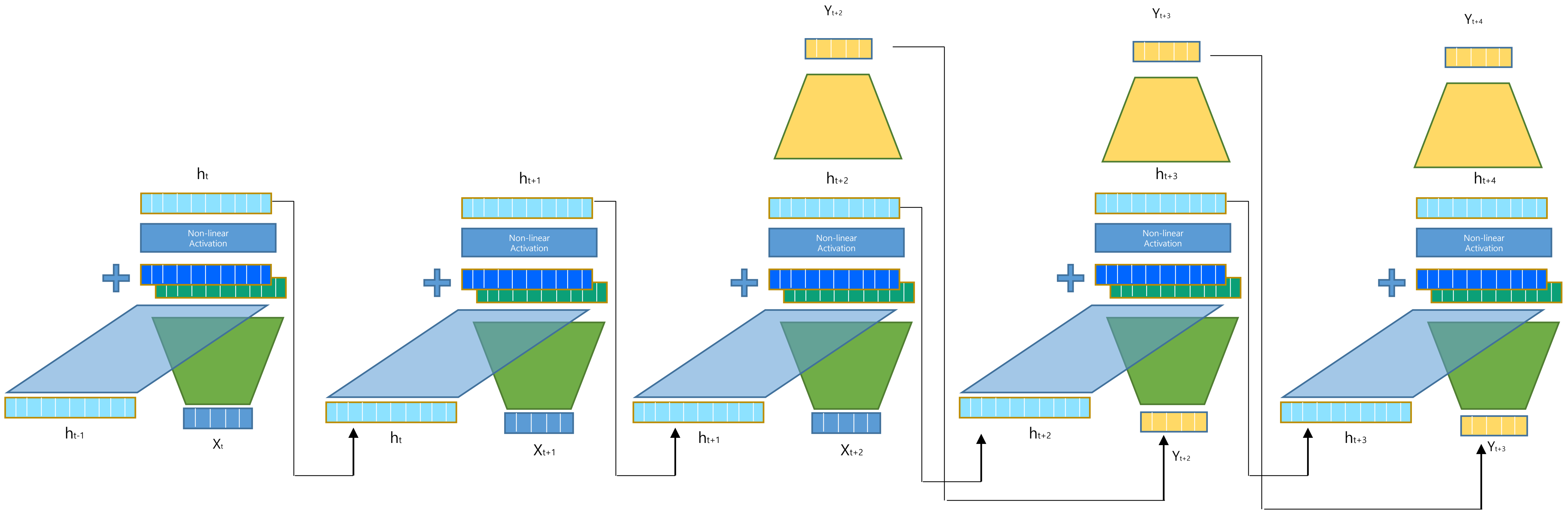




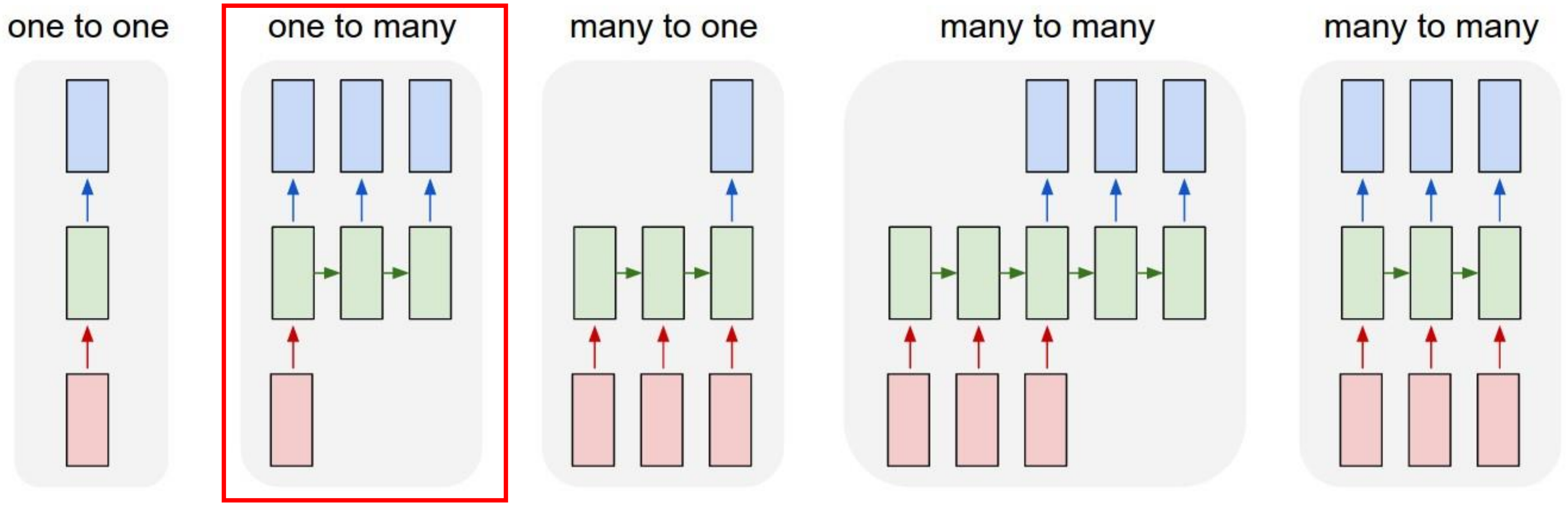
Recurrent Neural Network

many to many

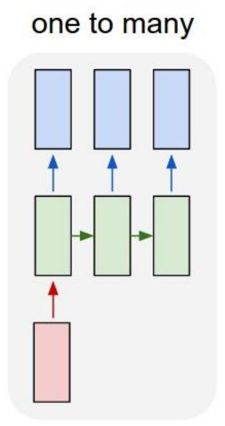


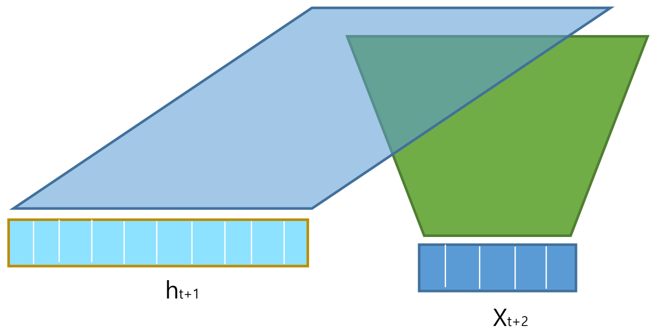


Types of Task Dealing with Sequential Data

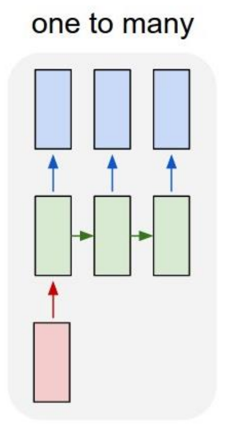


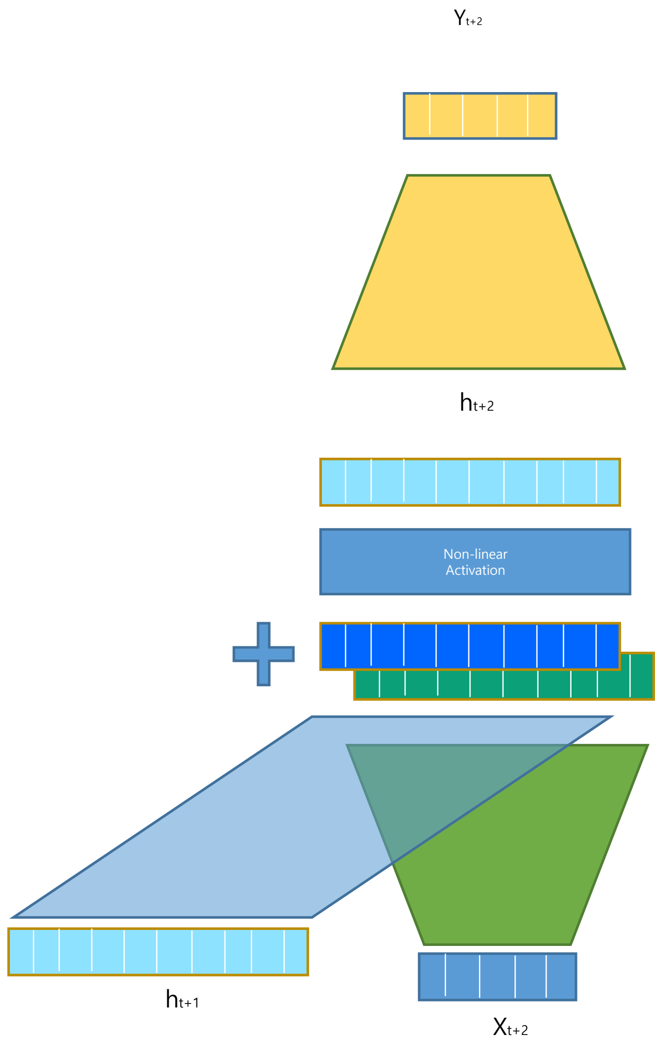
Recurrent Neural Network



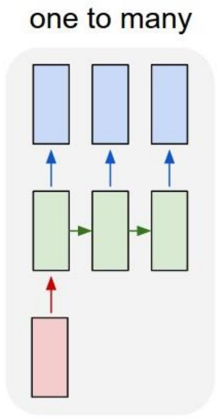


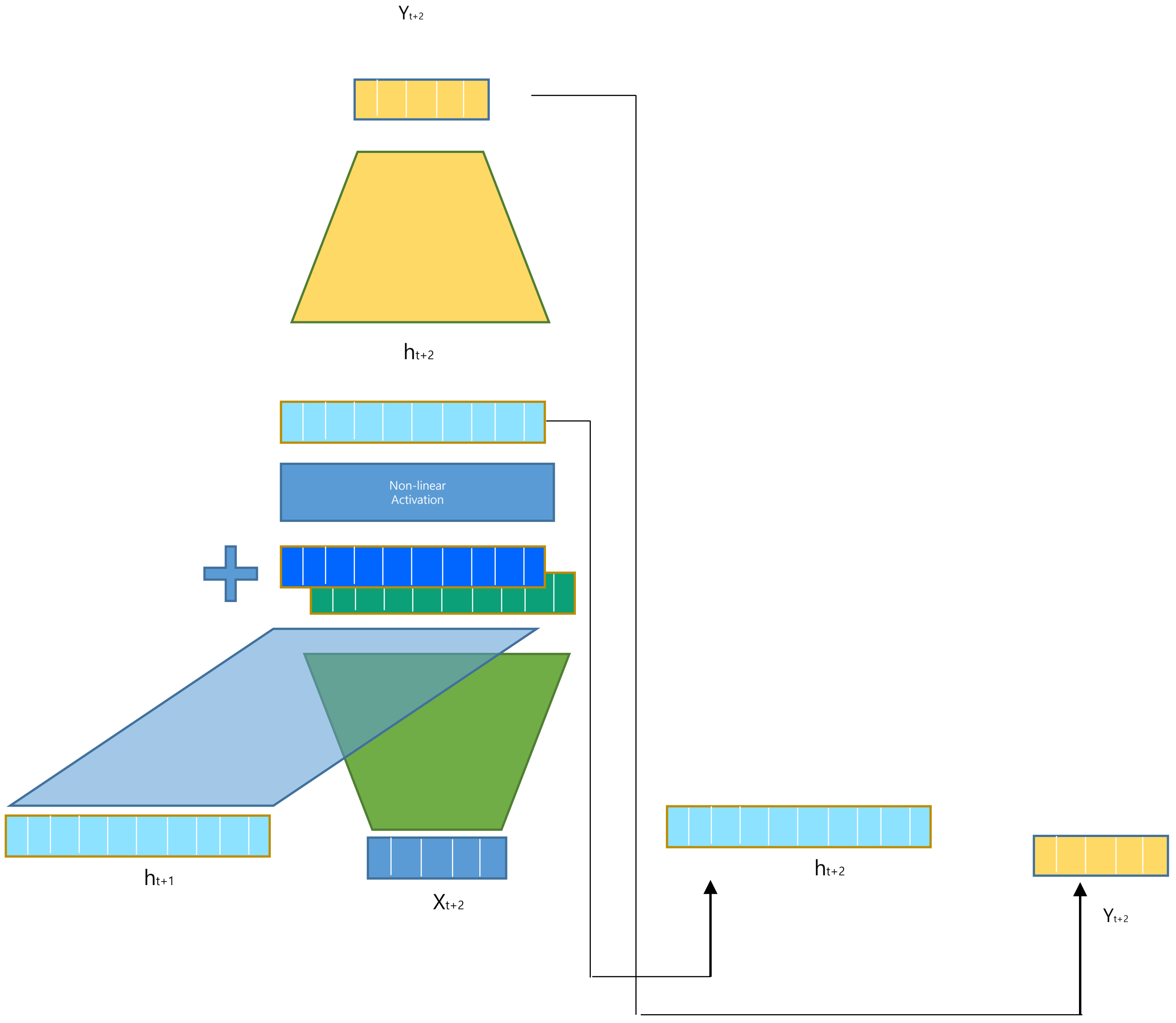
Recurrent Neural Network



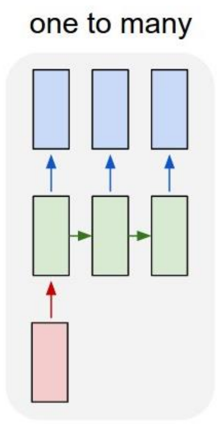


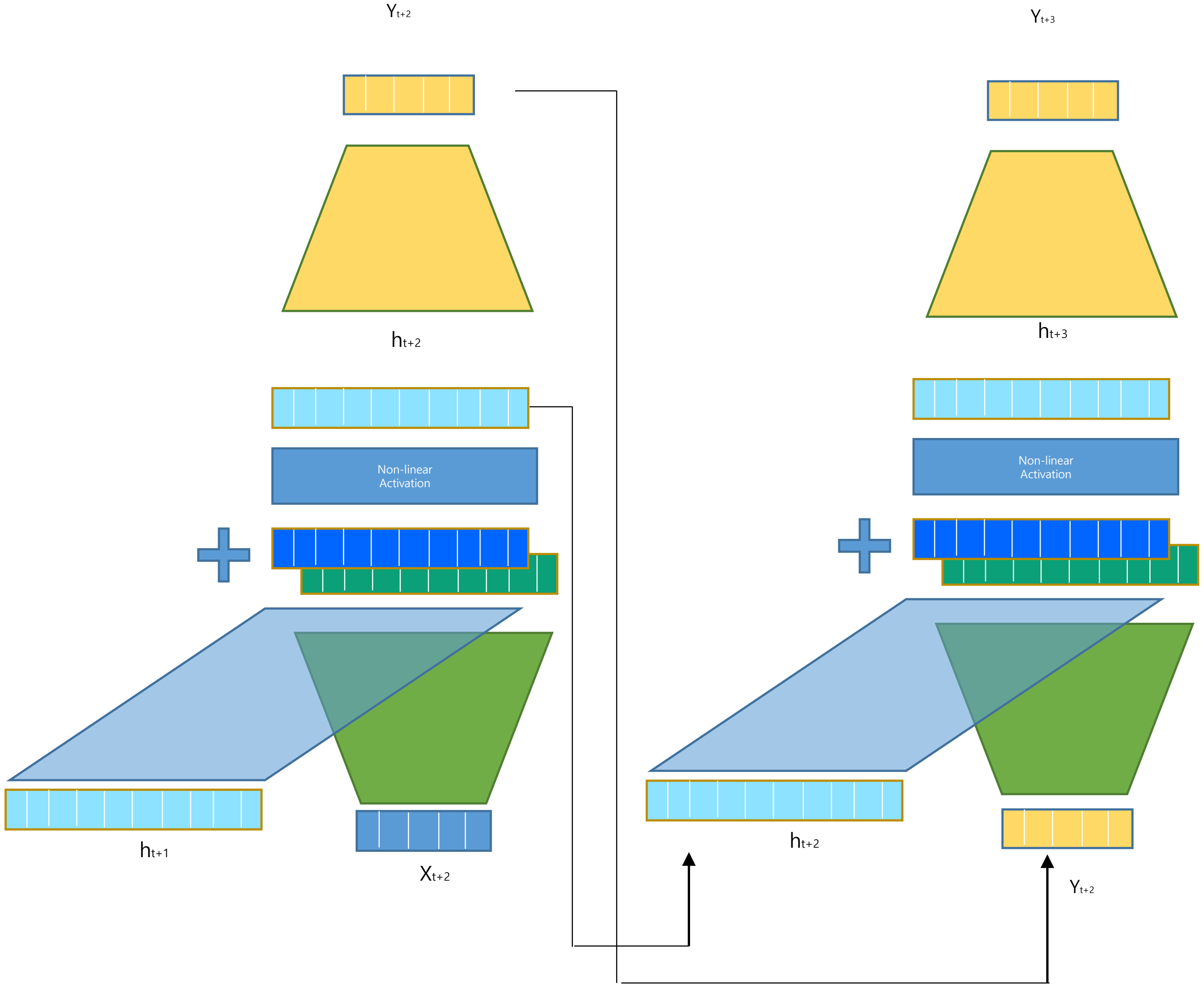
Recurrent Neural Network



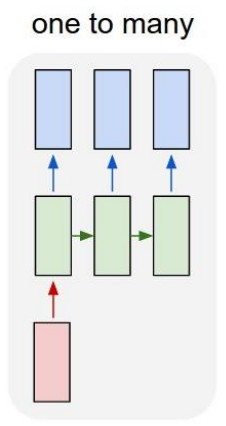


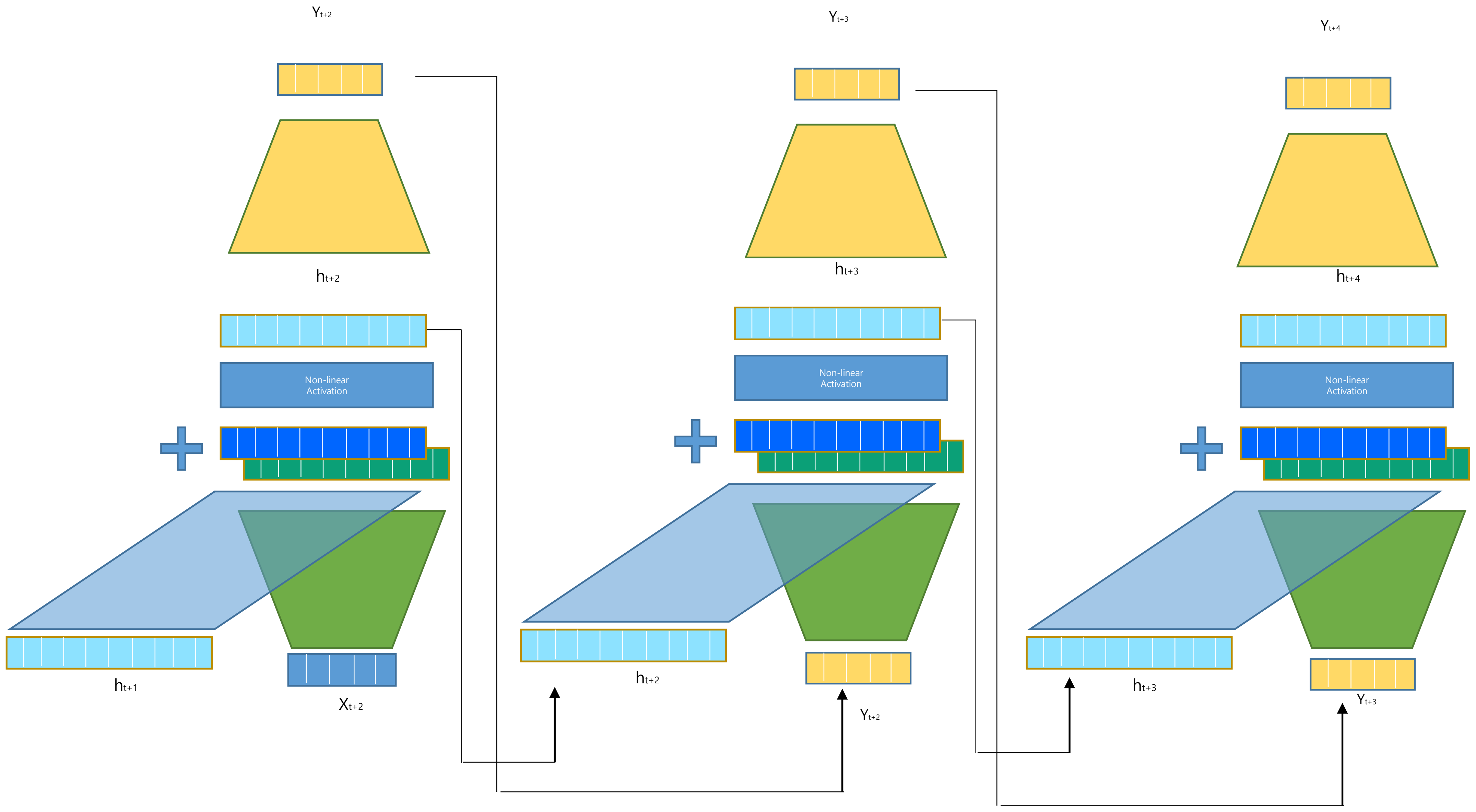
Recurrent Neural Network



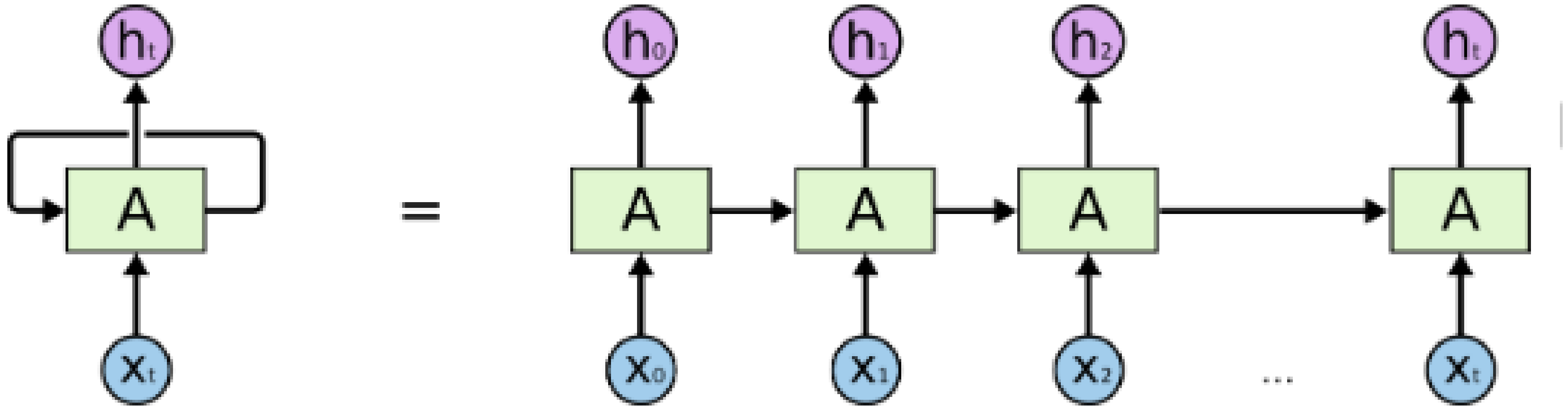


Recurrent Neural Network





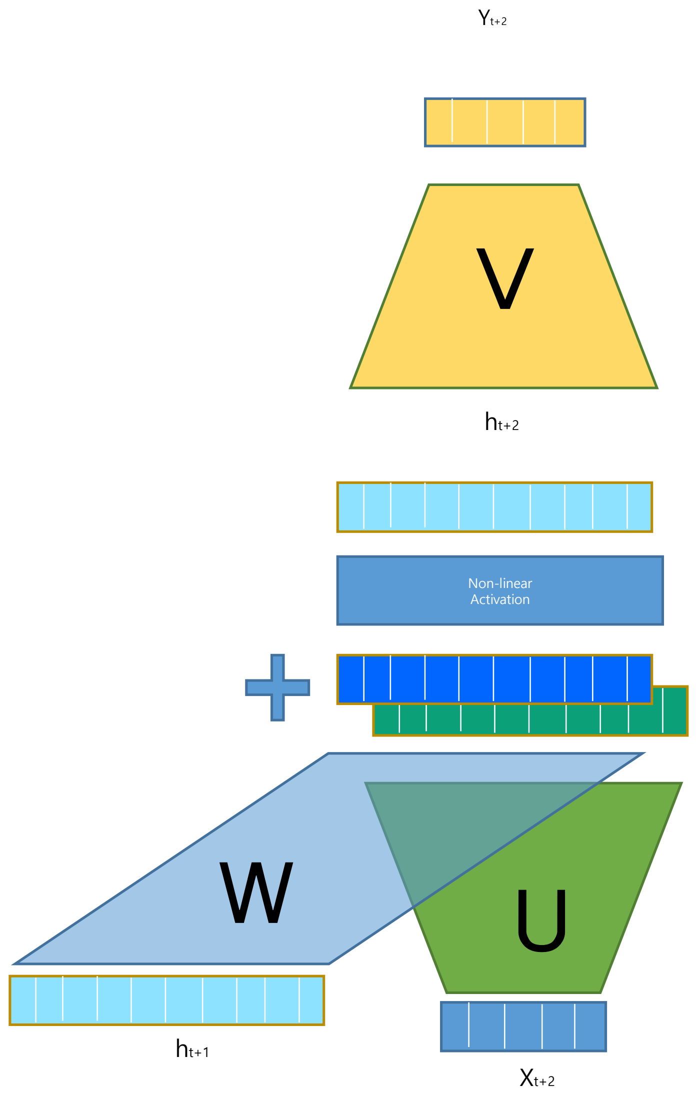
Recurrent Neural Network



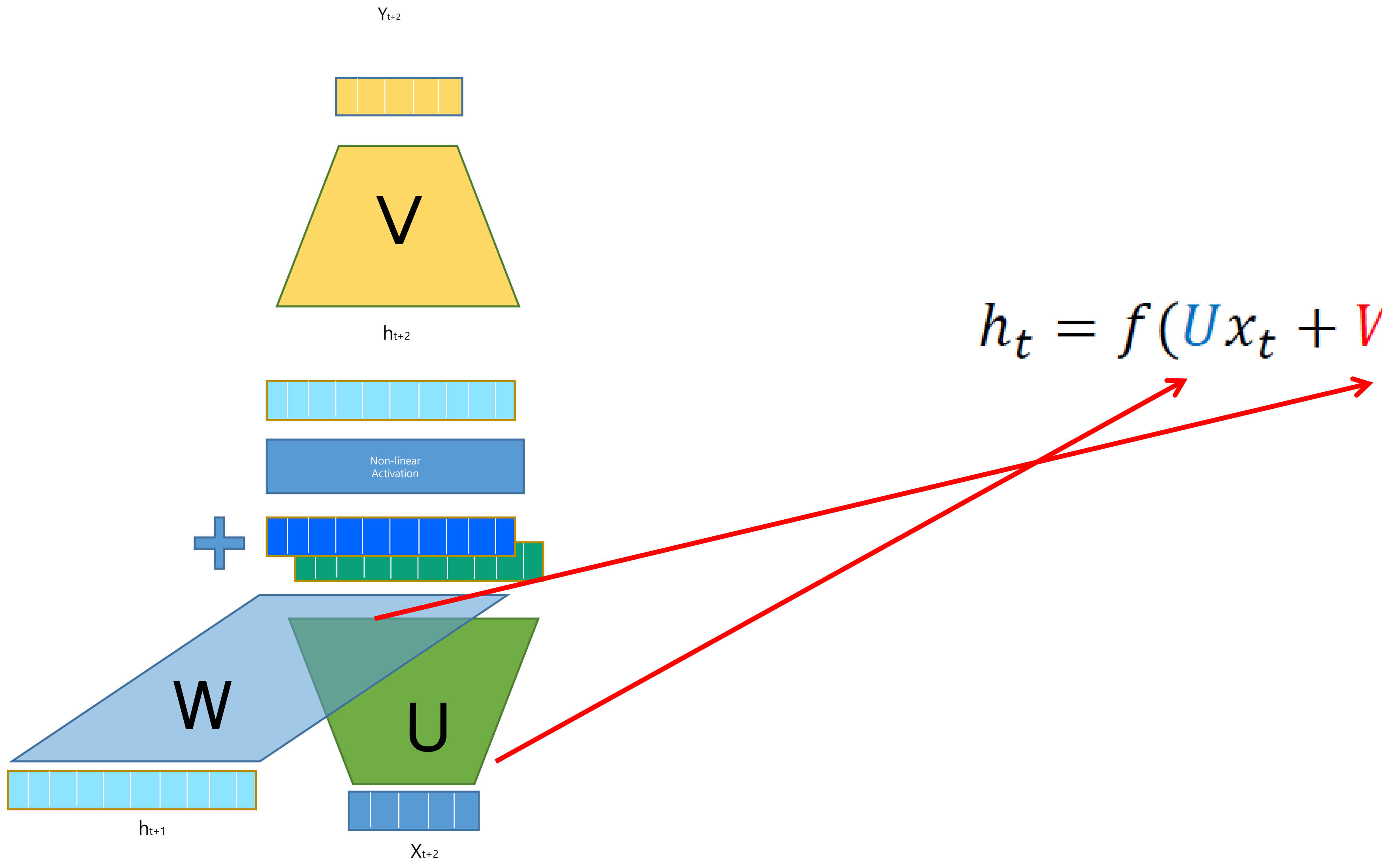
An unrolled recurrent neural network.

Recurrent Neural Network with Math

Recurrent Neural Network with Math

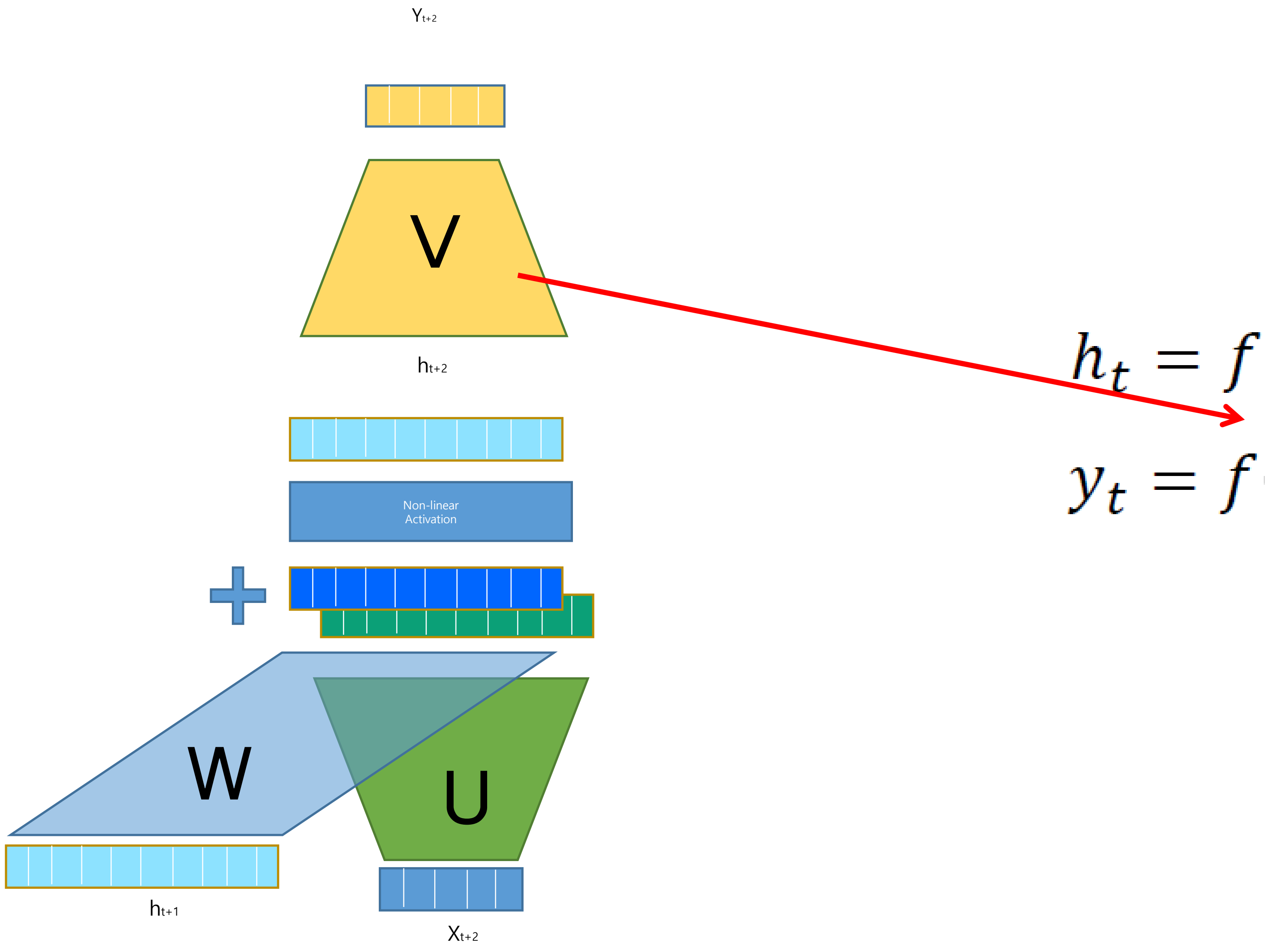


Recurrent Neural Network with Math



$$h_t = f(Ux_t + Wh_{t-1})$$

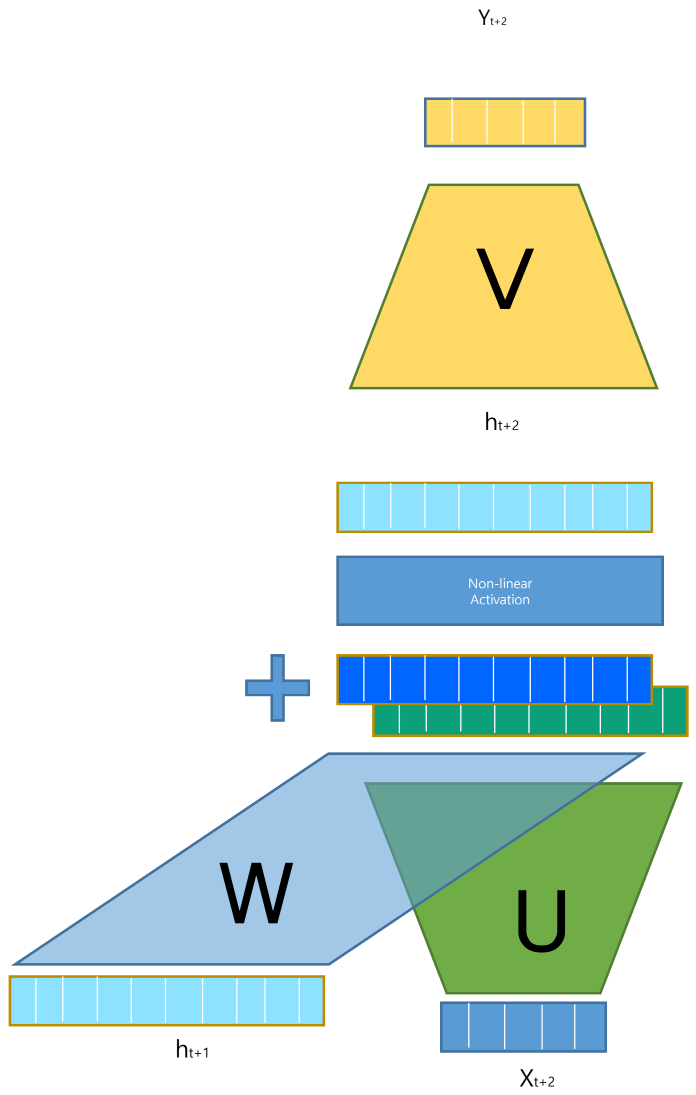
Recurrent Neural Network with Math



$$h_t = f(Ux_t + Wh_{t-1})$$

$$y_t = f(Vh_t)$$

Recurrent Neural Network with Math



f(x) = tanh(x)

$$h_t = f(Ux_t + Wh_{t-1})$$

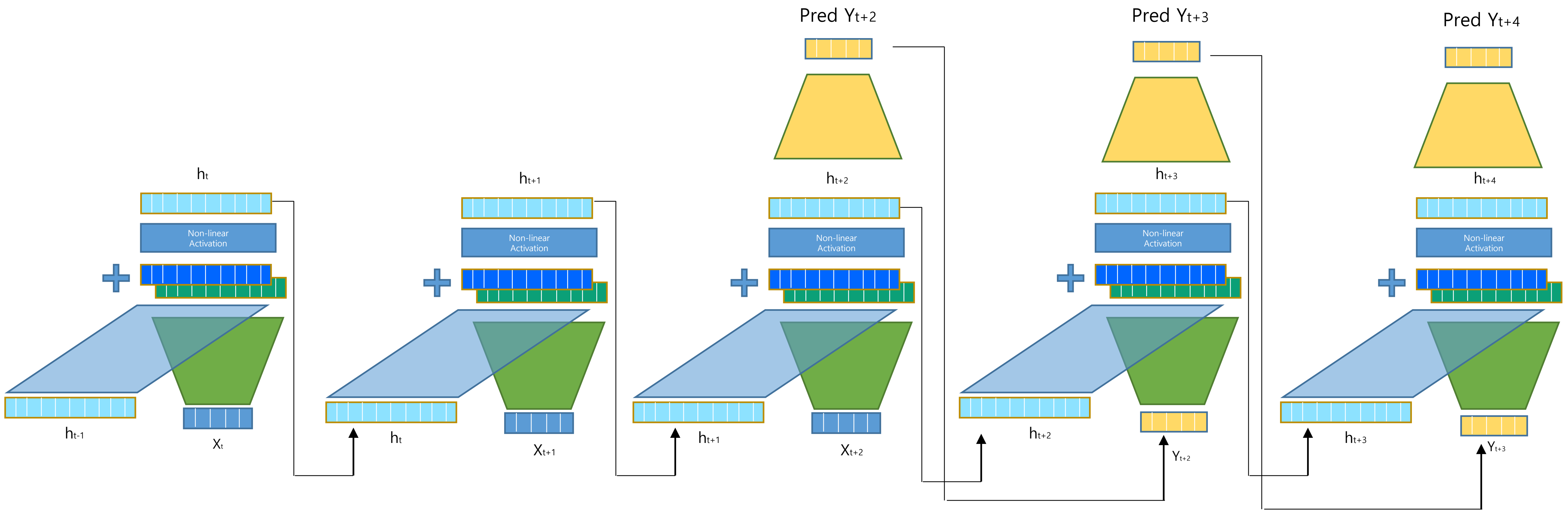
$$y_t = f(Vh_t)$$

f(x) = x

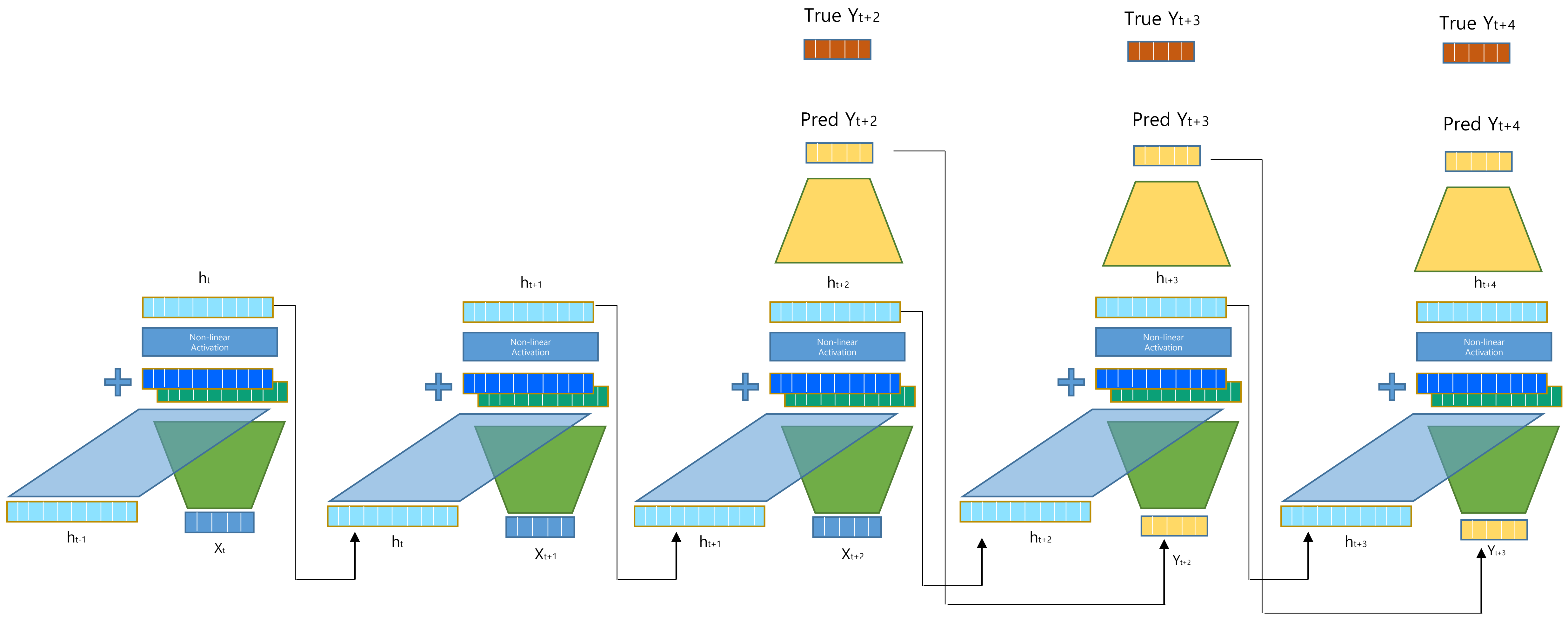
Okay, now we understand RNN model(hypothesis)

How can we evaluate it?

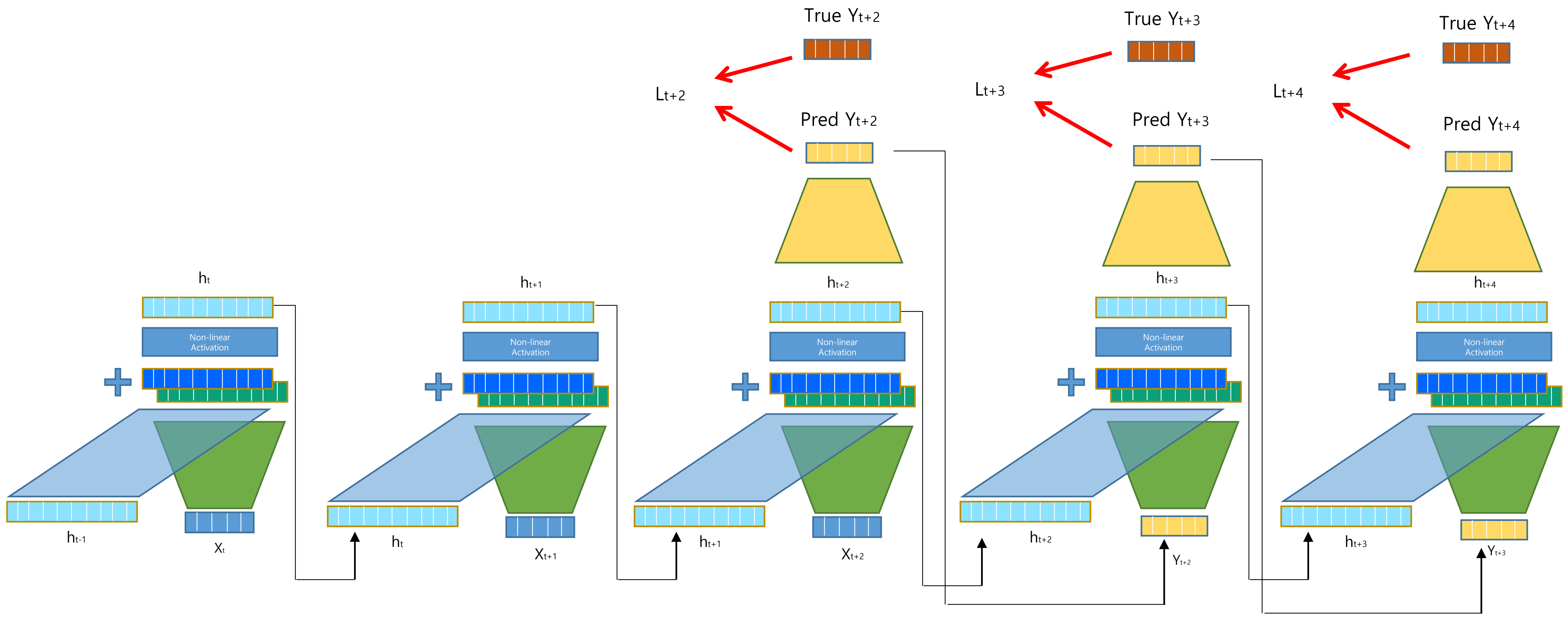
Calculate Loss of Recurrent Neural Network



Calculate Loss of Recurrent Neural Network

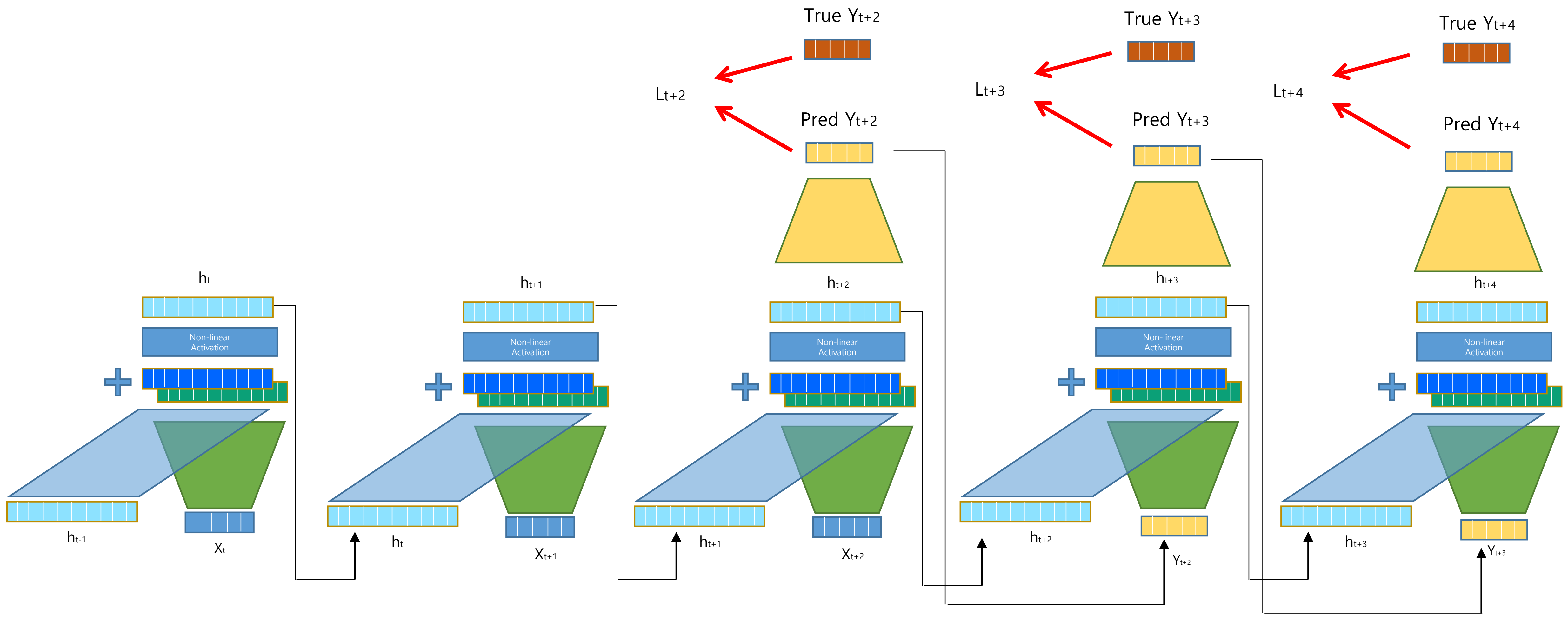


Calculate Loss of Recurrent Neural Network



Calculate Loss of Recurrent Neural Network

$$Loss(\theta) = \sum_t loss(y_{true,t}, y_{pred,t})$$

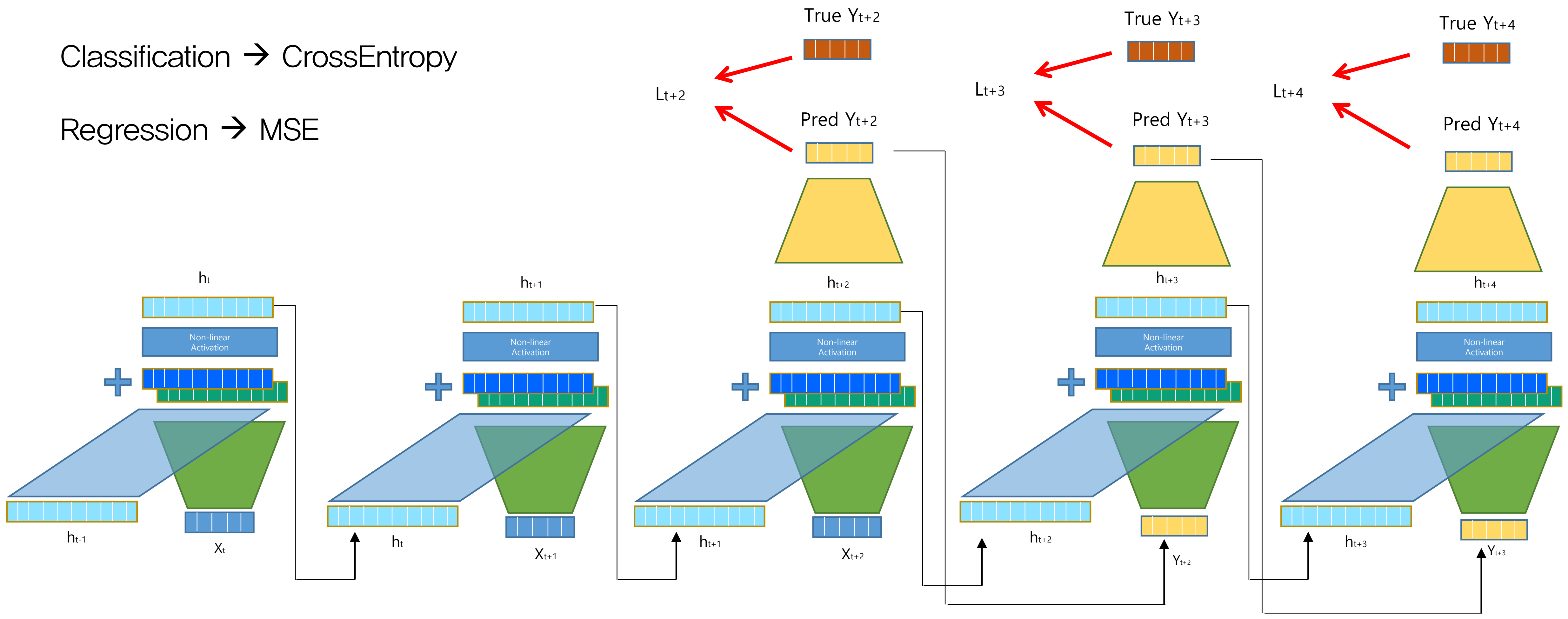


Calculate Loss of Recurrent Neural Network

$$Loss(\theta) = \sum_{t} loss(y_{true,t}, y_{pred,t})$$

Classification → CrossEntropy

Regression → MSE



Summary

- There are various tasks which have to deal with sequential data
- Recurrent Neural Network is suitable to handle it
- Basically RNN feeds new input and output from previous step together
- We can utilize RNN differently depends on the task

Today`s Time Schedule

Assignment #5 Review	20 mins
Recurrent Neural Network	1 hour
Implement Basic RNN in Pytorch	1.5 hour