Deep Learning

Vazgen Mikayelyan

YSU, Krisp

October 31, 2019

Outline

1. Metrics for Classification Problems

2. Basic Recurrent Neural Networks

ROC Curve

ROC Curve

- A receiver operating characteristic curve, i.e. ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.

ROC Curve

- A receiver operating characteristic curve, i.e. ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.
- The diagnostic performance of a test or the accuracy of a test to discriminate diseased cases from normal cases is evaluated using ROC curve analysis.

ROC Curve

- A receiver operating characteristic curve, i.e. ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.
- The diagnostic performance of a test or the accuracy of a test to discriminate diseased cases from normal cases is evaluated using ROC curve analysis.
- A ROC curve is a way to compare diagnostic tests. It is a plot of true positive rate against the false positive rate:

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

- What if $threshold = 0$?

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

- What if $threshold = 0$?
  - $TN = 0, FN = 0$

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

- What if $threshold = 0$?
  - $TN = 0, FN = 0$
  - $TPR = 1, FPR = 1$

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

- What if $threshold = 0$?
  - $TN = 0, FN = 0$
  - $TPR = 1, FPR = 1$
- What if $threshold = 1$?

ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

- What if $threshold = 0$?
  - $TN = 0, FN = 0$
  - $TPR = 1, FPR = 1$
- What if $threshold = 1$?
  - $TP = 0, FP = 0$

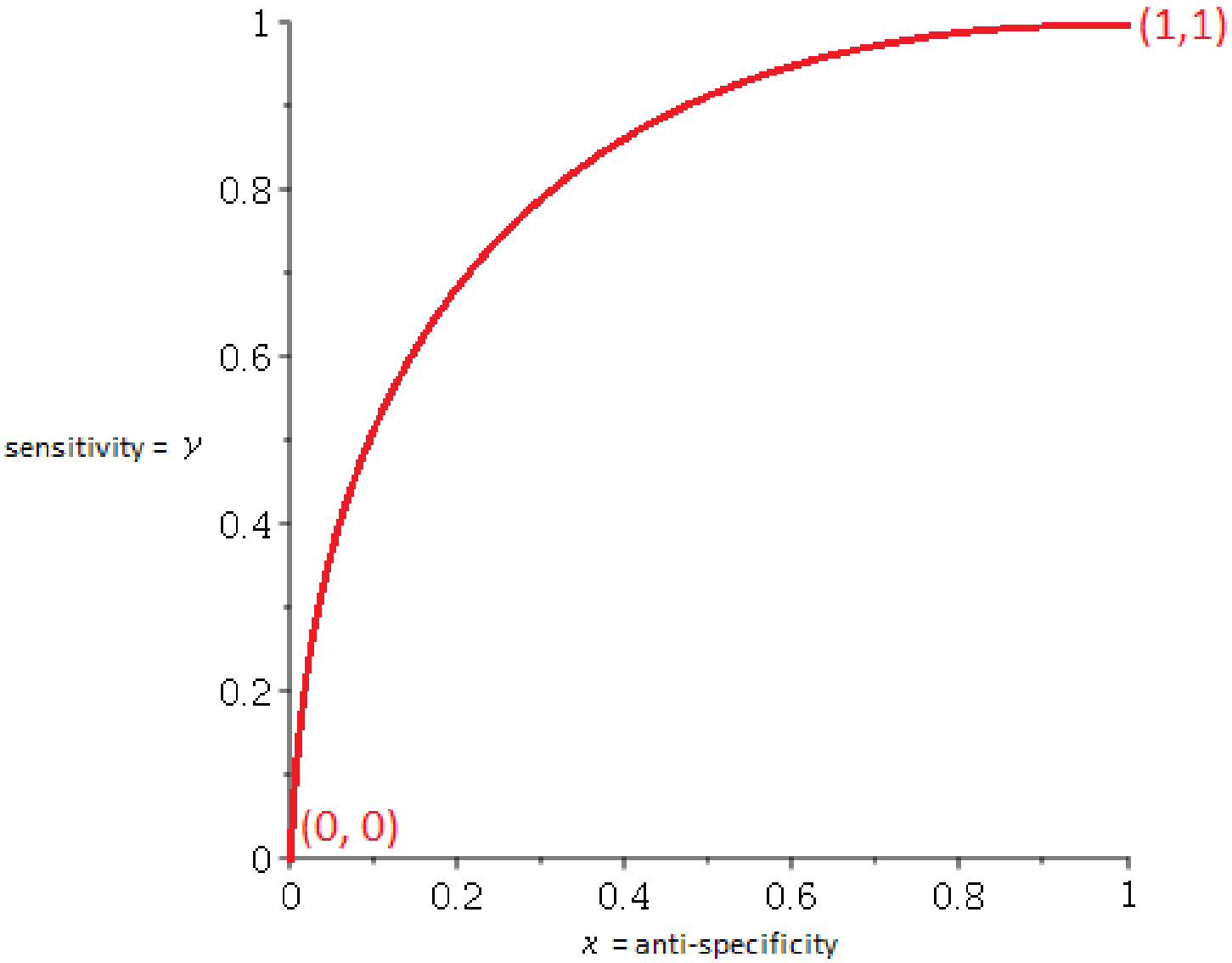
ROC Curve

$$TPR = Recall = Sensitivity = \frac{TP}{TP + FN}$$

$$FPR = 1 - Specificity = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

- What if $threshold = 0$?
  - $TN = 0, FN = 0$
  - $TPR = 1, FPR = 1$
- What if $threshold = 1$?
  - $TP = 0, FP = 0$
  - $TPR = 0, FPR = 0$

ROC Curve


1
0.8
0.6
0.4
0.2
0
sensitivity = y
(1,1)
(0, 0)
0
0.2
0.4
0.6
0.8
1
x = anti-specificity

ROC Curve

Definition 1

AUC (area under the curve) is the area under the ROC curve.

ROC Curve

Definition 1

AUC (area under the curve) is the area under the ROC curve.

- What if $AUC \approx 0.5$?

ROC Curve

Definition 1

AUC (area under the curve) is the area under the ROC curve.

- What if $AUC \approx 0.5$?
- What if $AUC \approx 1$?

ROC Curve

Definition 1

AUC (area under the curve) is the area under the ROC curve.

- What if $AUC \approx 0.5$?
- What if $AUC \approx 1$?
- How to choose threshold?

ROC Curve

Definition 1

AUC (area under the curve) is the area under the ROC curve.

- What if $AUC \approx 0.5$?
- What if $AUC \approx 1$?
- How to choose threshold?
- How to use ROC curve for multi-class model?

Balanced Cross Entropy

- Simple Cross Entropy

$$L(w) = \frac{1}{n} \sum_{i=1}^{n} \left(-y_i^T \log f_w(x_i) \right).$$

Balanced Cross Entropy

- Simple Cross Entropy

$$L(w) = \frac{1}{n} \sum_{i=1}^{n} \left(-y_i^T \log f_w(x_i) \right).$$

- Balanced Cross Entropy

$$L(w) = \frac{1}{n} \sum_{i=1}^{n} \left(-(\alpha * y_i)^T \log f_w(x_i) \right).$$

Where $\alpha^T = [\alpha_1, \ldots, \alpha_k]$, $\alpha_i \in (0, 1)$ and $\sum_{i=1}^{k} \alpha_i = 1$.

Outline

1 Metrics for Classification Problems

2 Basic Recurrent Neural Networks

Sequential Data

- What is sequential data?

Sequential Data

- What is sequential data?
  - Time series

Sequential Data

- What is sequential data?
  - Time series
  - Text

Sequential Data

- What is sequential data?
  - Time series
  - Text
  - Speech

Sequential Data

- What is sequential data?
  - Time series
  - Text
  - Speech
  - Audio

Sequential Data

- What is sequential data?
  - Time series
  - Text
  - Speech
  - Audio
  - Video

Sequential Data

- What is sequential data?
  - Time series
  - Text
  - Speech
  - Audio
  - Video
- Why don't to use known networks?

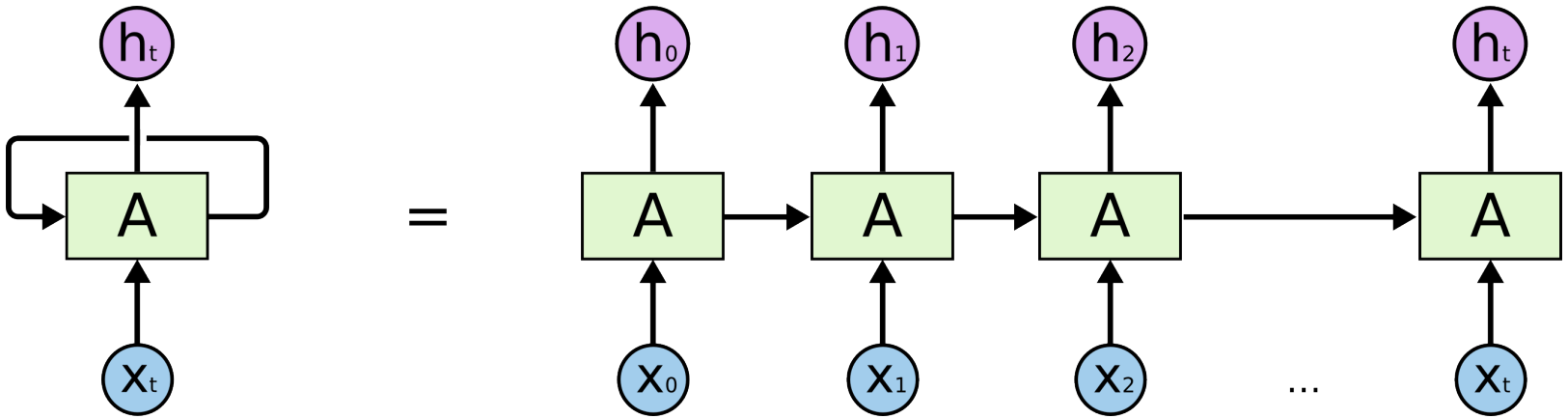
Sequential Data

- What is sequential data?
  - Time series
  - Text
  - Speech
  - Audio
  - Video
- Why don't to use known networks?
  - Inputs, outputs can be different lengths in different examples.

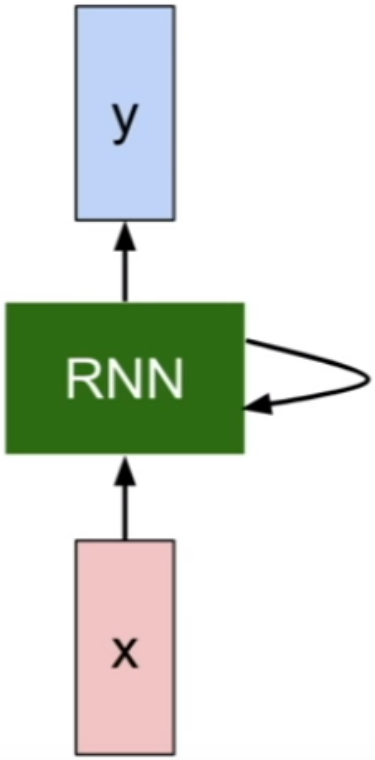
Sequential Data

- What is sequential data?
  - Time series
  - Text
  - Speech
  - Audio
  - Video
- Why don't to use known networks?
  - Inputs, outputs can be different lengths in different examples.
  - Doesn't share features learned across different positions of sequence.

Basic RNNs



Basic RNNs



$$h_t = f_W(h_{t-1}, x_t)$$

$$\downarrow$$

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

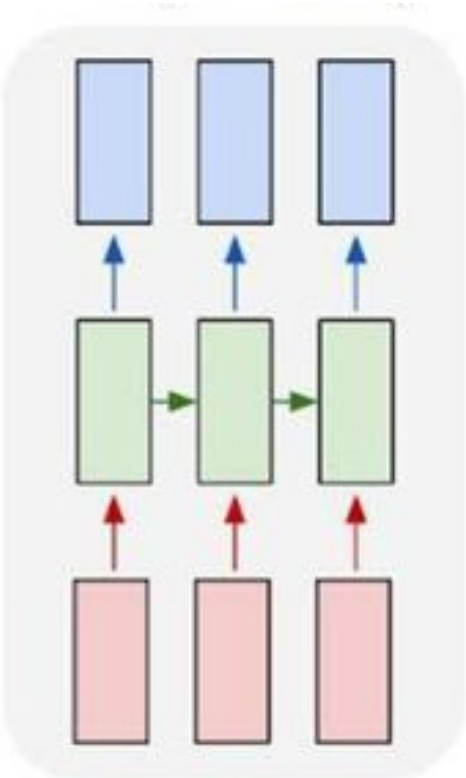
$$y_t = W_{hy}h_t$$

Different Types of RNNs

- Many to many

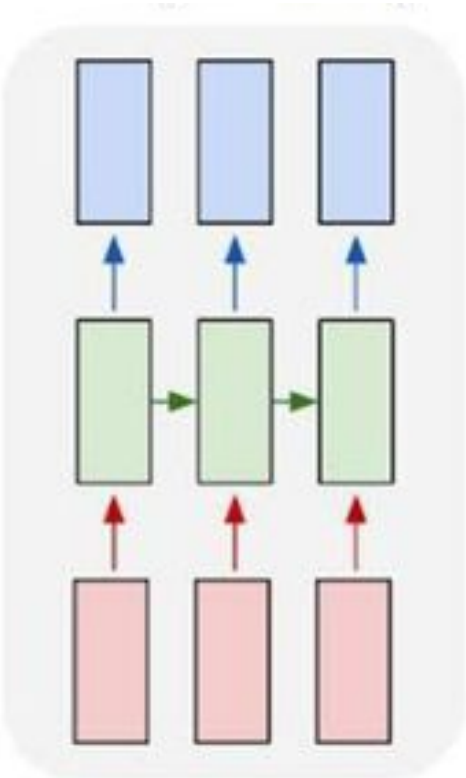
Different Types of RNNs

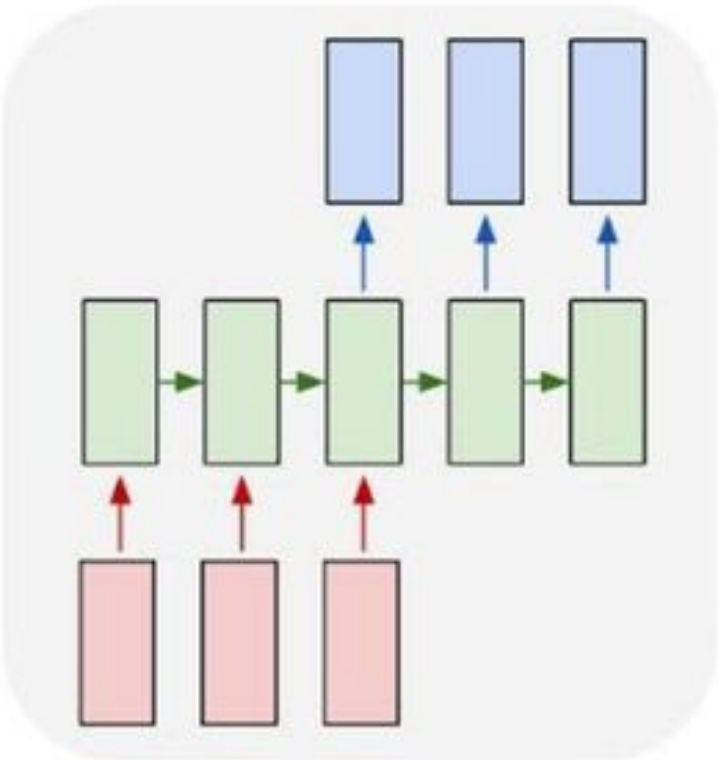
- Many to many



Different Types of RNNs

- Many to many



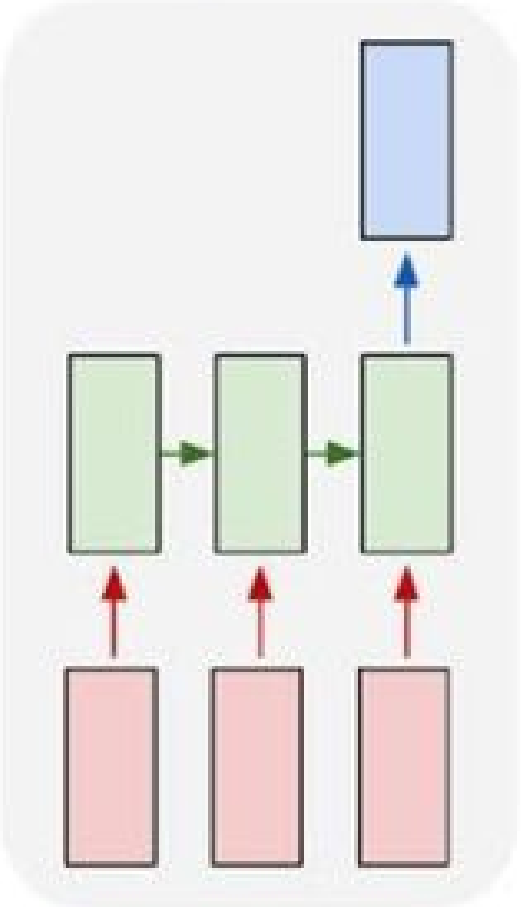


Different Types of RNNs

- Many to one

Different Types of RNNs

- Many to one

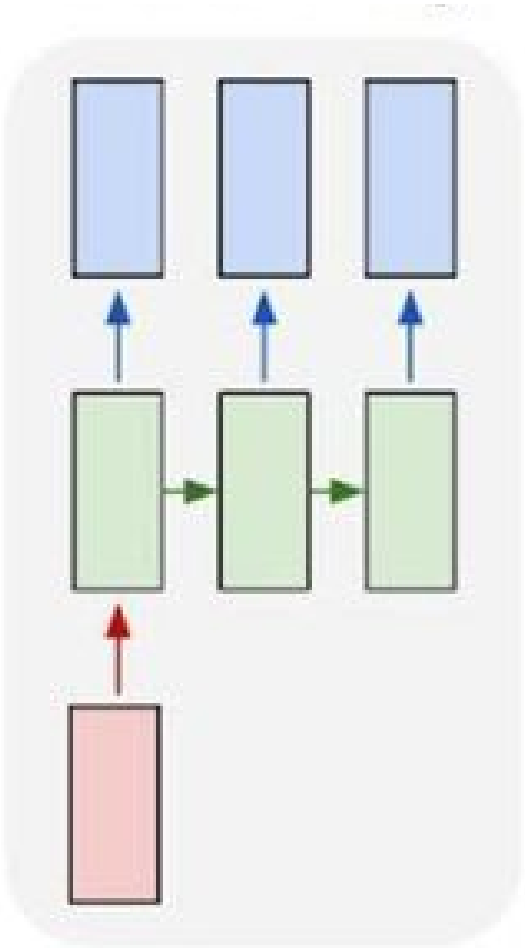


Different Types of RNNs

- One to many

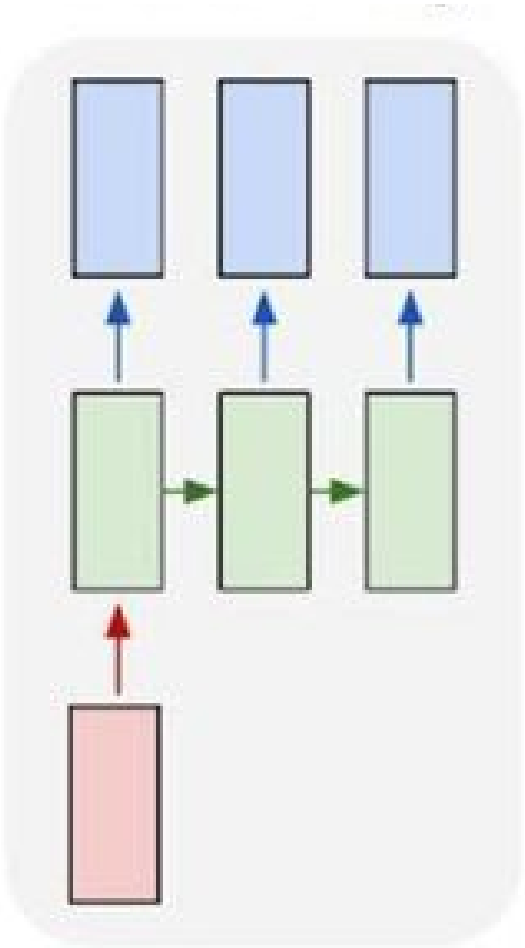
Different Types of RNNs

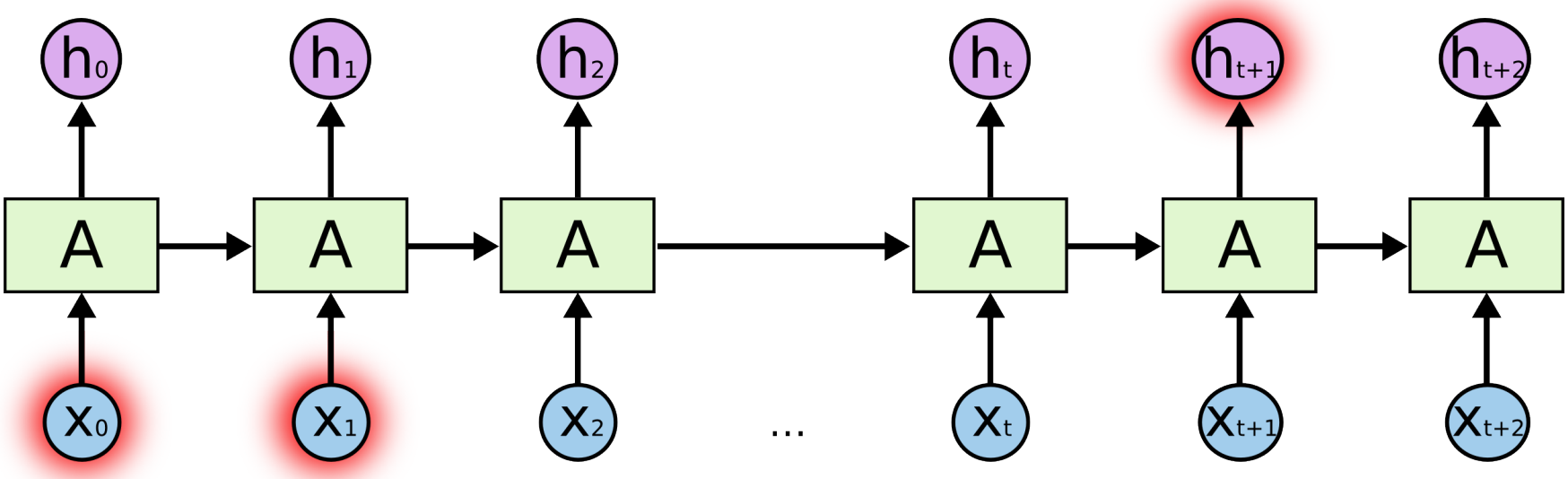
- One to many



Different Types of RNNs

- One to many



- One to one?

Problem of Long Term Dependencies