01 TECART | SMART TECHNOLOGY

RNN

The team

Nara Surya (Presenter)

Ketua Divisi Smart Technology Tecart

Jana

Anggota Divisi Smart Technology Tecart

Gede Ocha

Anggota Divisi Smart Technology Tecart

Recap Workshop week 2

1. Proses Konvolusi
2. Convolutional Neural Network
3. Sedikit Computer Vision
4. Mengklasifikasi Wasp and Bee

Materi Hari ini

1. Apa itu Recurrent Neural Network ?
2. Permasalahan-permasalahan yang umum terjadi di RNN
3. architecture RNN
4. Mengklasifikasi Sentiment

Available :

On Recording

https://github.com/NaraSurya/Deep-Learning-Workshop-Tecart-2020

Pengenalan RNN

Apa itu RNN ?

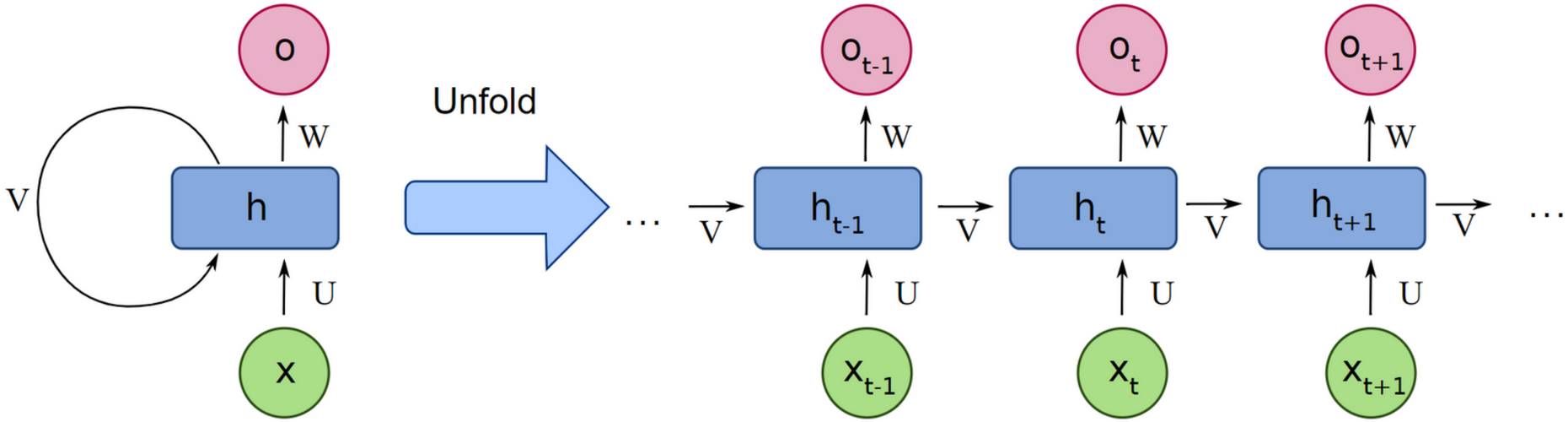
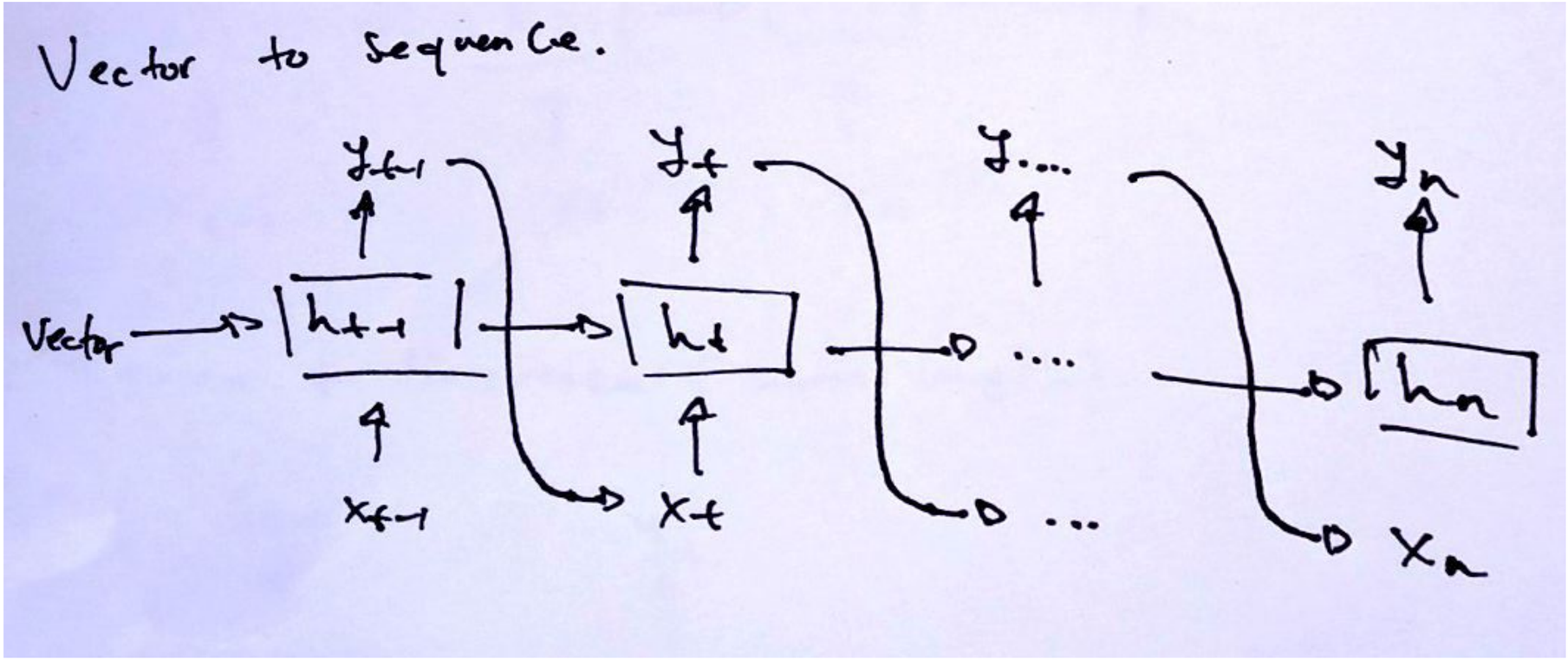


Image By https://medium.com/deeplearningbrasilia/deep-learning-recurrent-neural-networks-f9482a24d010

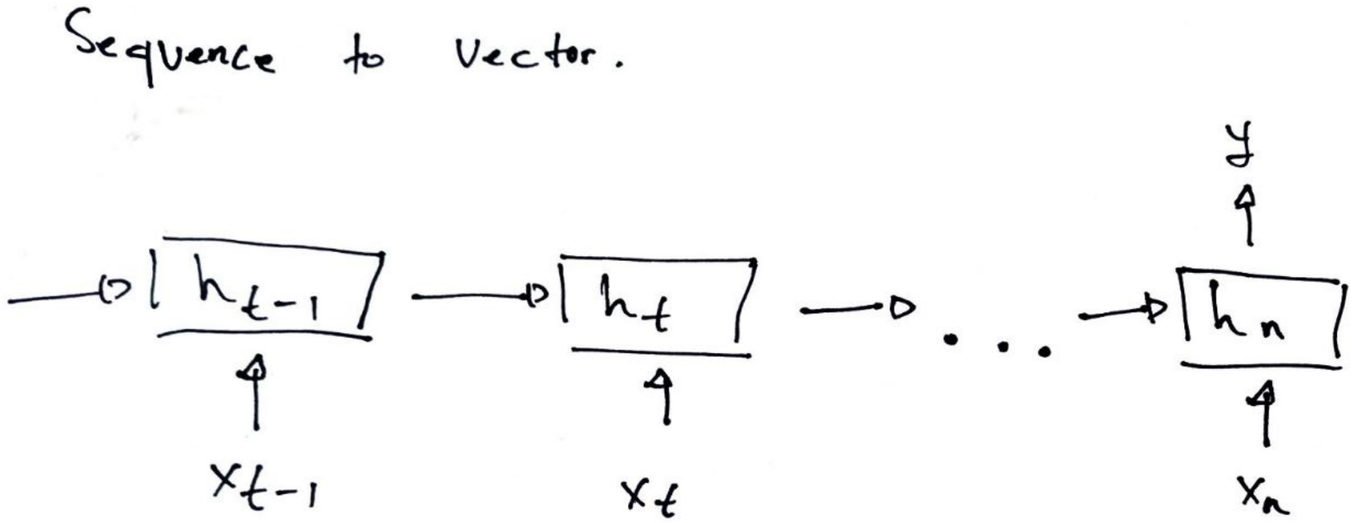
Generative RNN



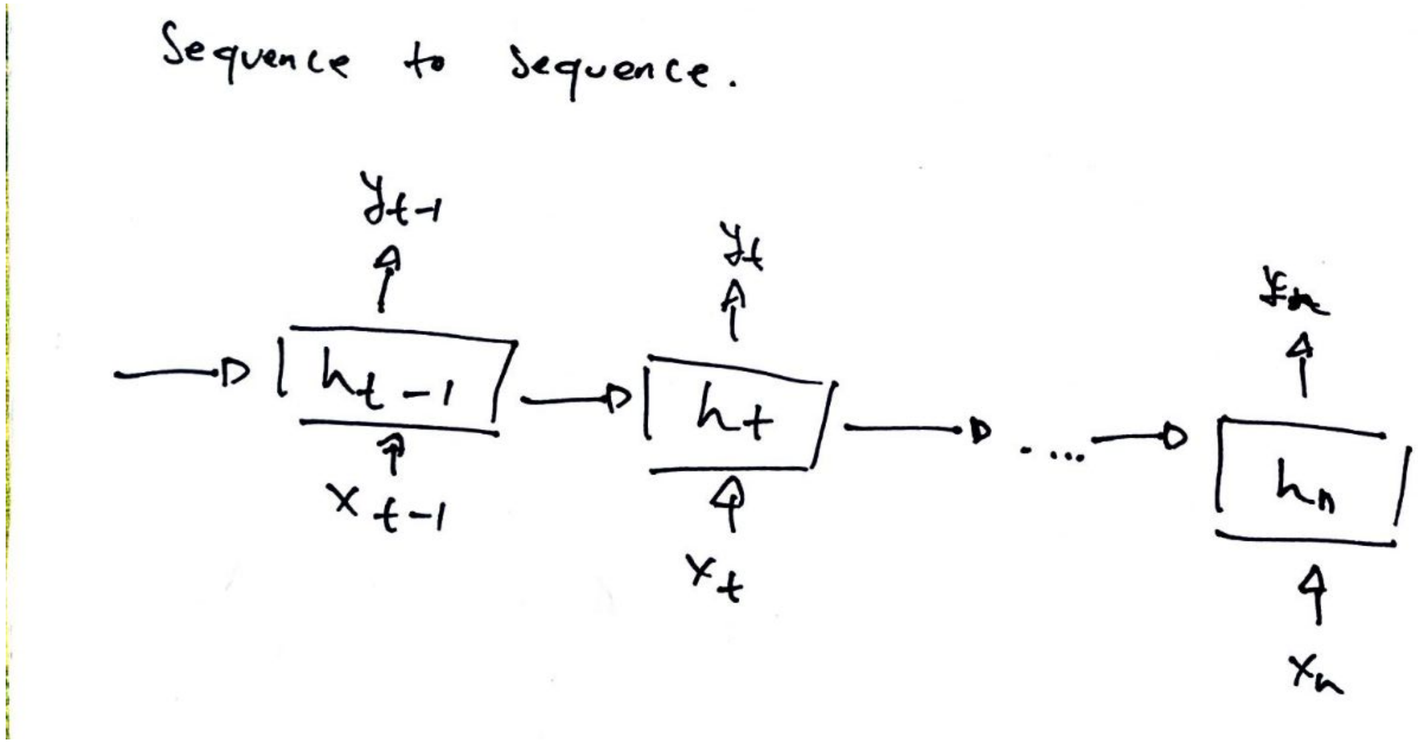
conditional distribution

1. Sequence to Vector
2. Sequence to Sequence (same length)
3. Vector to Sequence
4. Sequence to Sequence

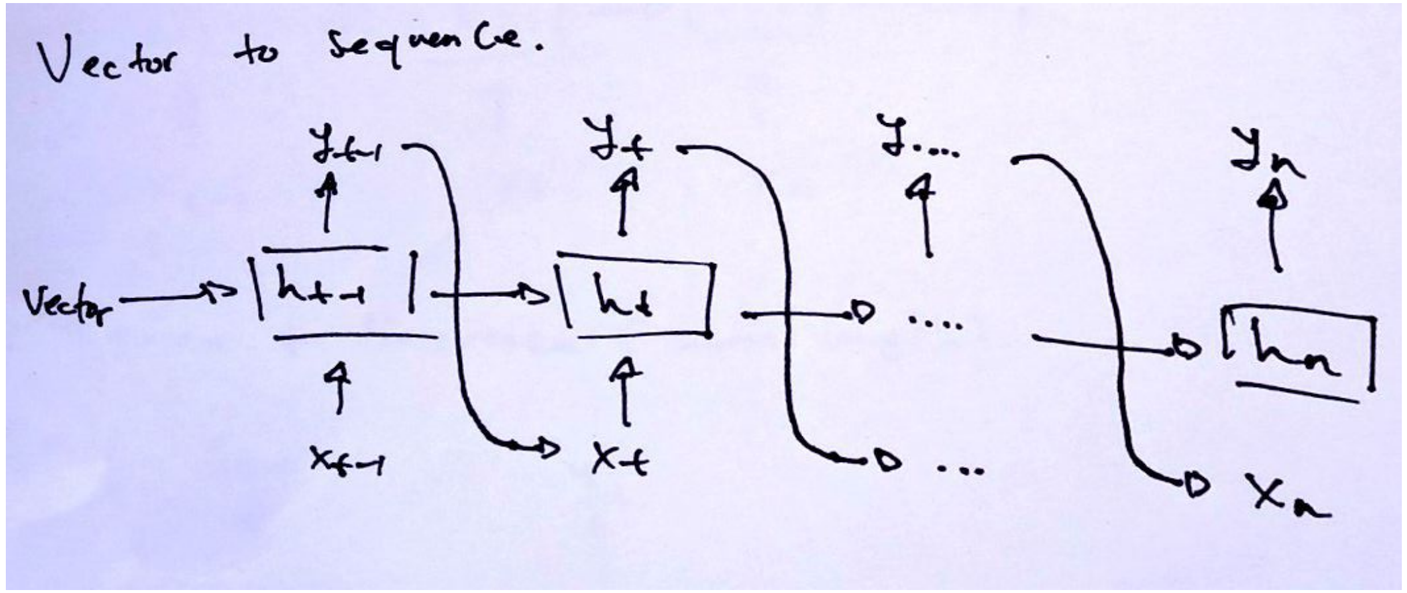
Sequence to vector



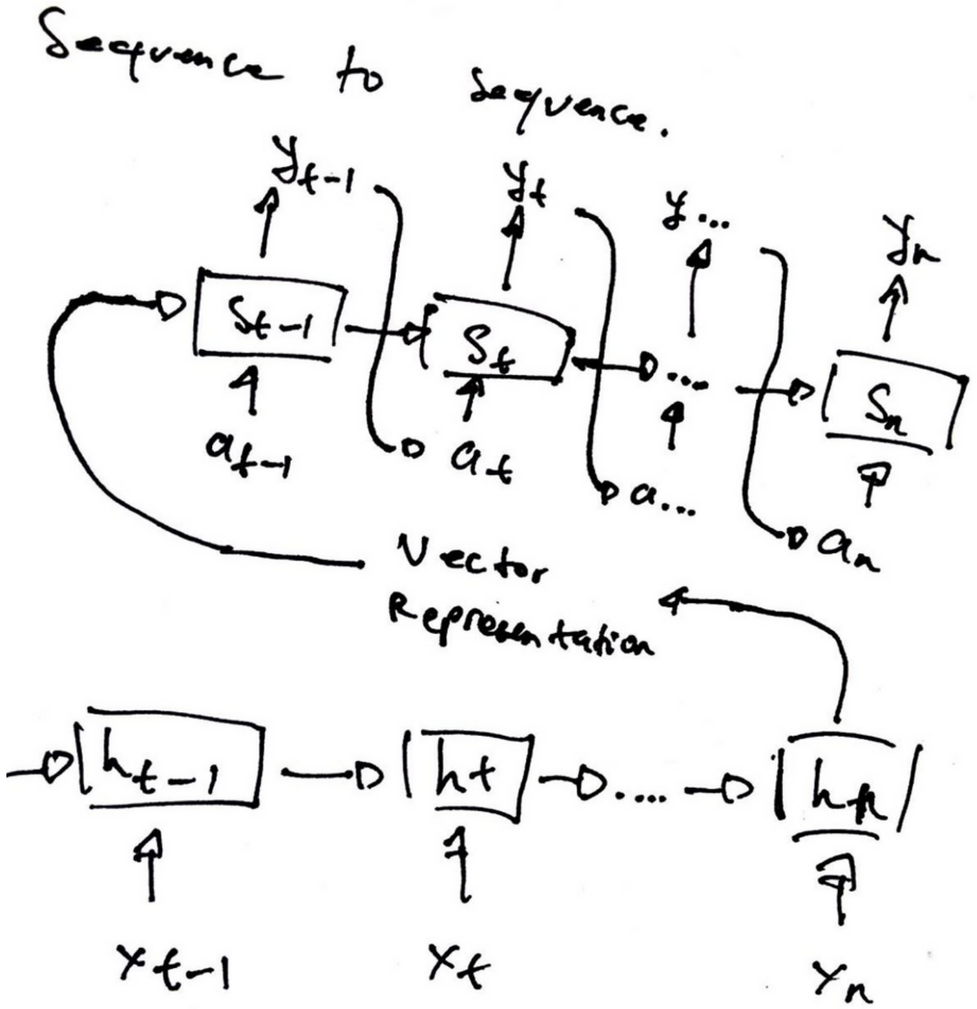
Sequence to Sequence (Same length)



Vector to Sequence



Sequence to Sequence



Permasalahan dalam RNN

Backpropagation Through Time

$$\frac{\partial E_3}{\partial W} = \sum_{k=0}^{3} \frac{\partial E_3}{\partial \hat{y}_3} \frac{\partial \hat{y}_3}{\partial s_3} \frac{\partial s_3}{\partial s_k} \frac{\partial s_k}{\partial W}$$

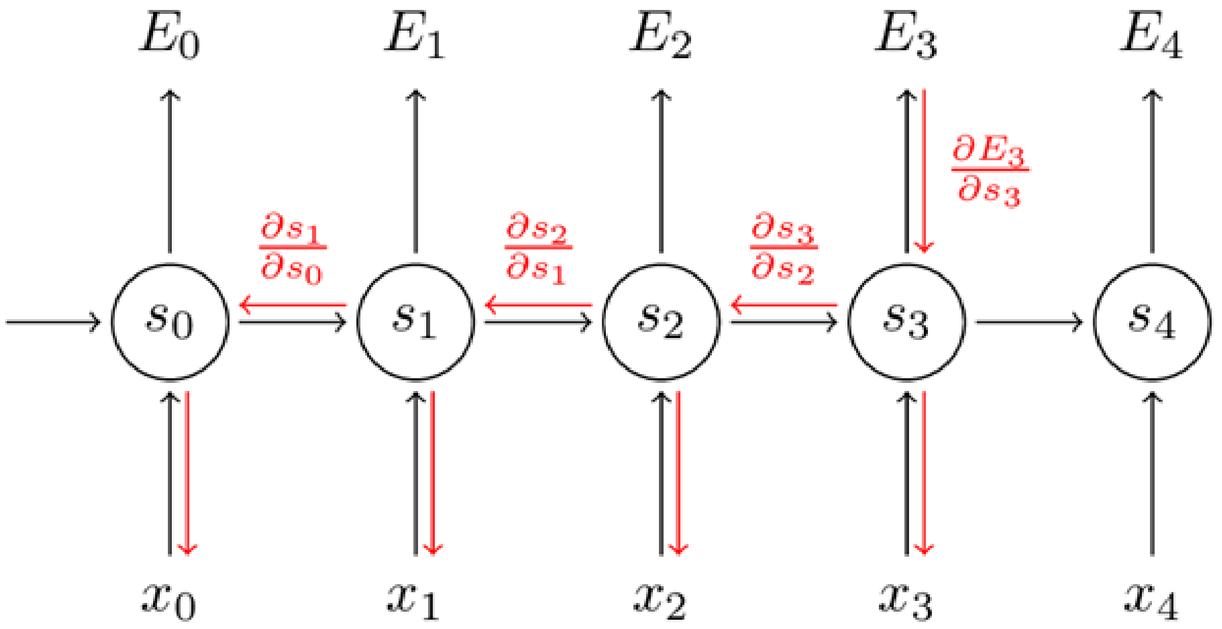
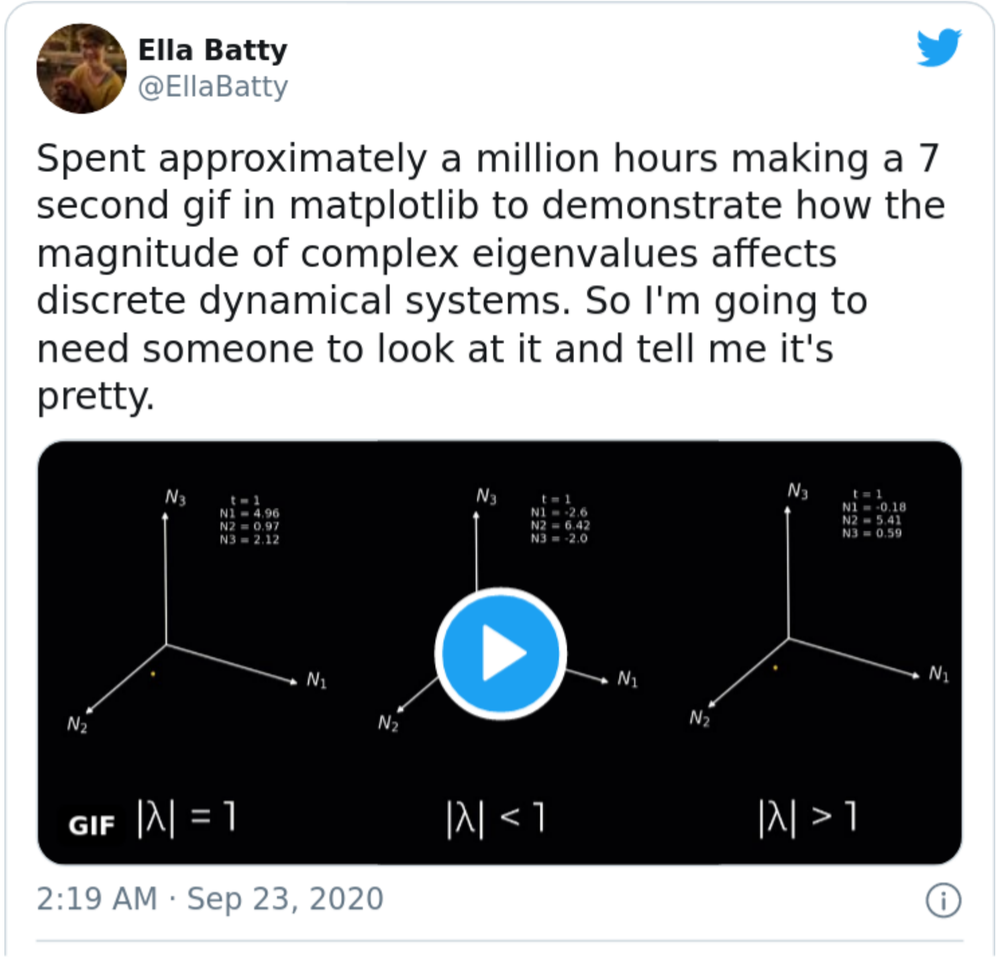


Image By https://medium.com/@madhuramiah/bi-directional-rnn-basics-of-lstm-and-gru-e114aa4779bb

Vanishing or Exploding Gradient Problem



Gradient Clipping

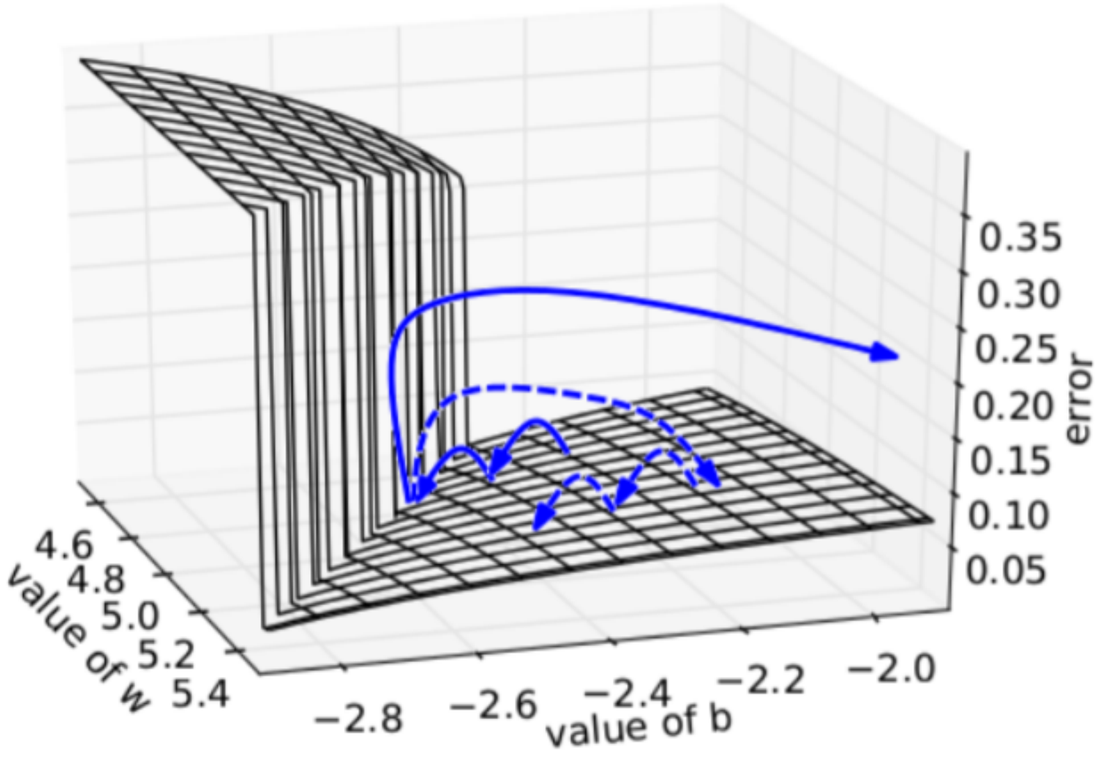


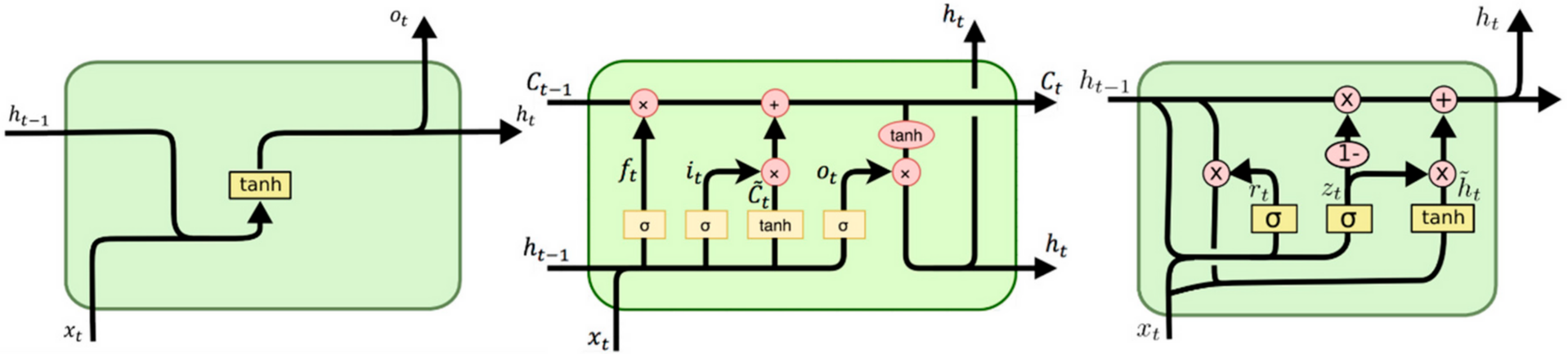
Image By Pascanu, Razvan et al. “On the difficulty of training recurrent neural networks.” ICML (2013).

Algorithm 1 Pseudo-code for norm clipping the gradients whenever they explode

$\hat{\mathbf{g}} \leftarrow \frac{\partial \mathcal{E}}{\partial \theta}$
if $\|\hat{\mathbf{g}}\| \geq threshold$ **then**
  $\hat{\mathbf{g}} \leftarrow \frac{threshold}{\|\hat{\mathbf{g}}\|}\hat{\mathbf{g}}$
end if

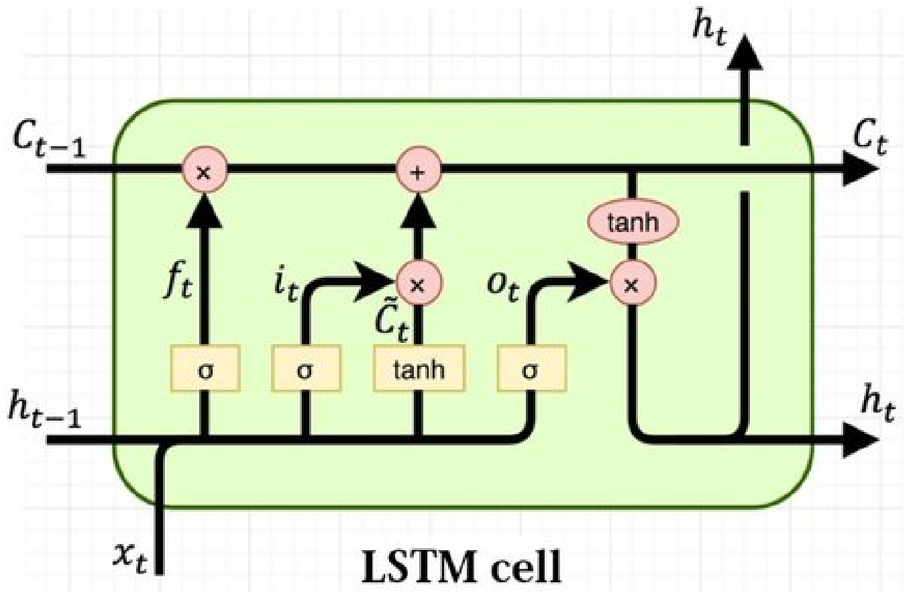
Learn More : Pascanu, Razvan et al. “On the difficulty of training recurrent neural networks.” ICML (2013).

LSTM dan GRU



Simple RNN **LSTM** **GRU**

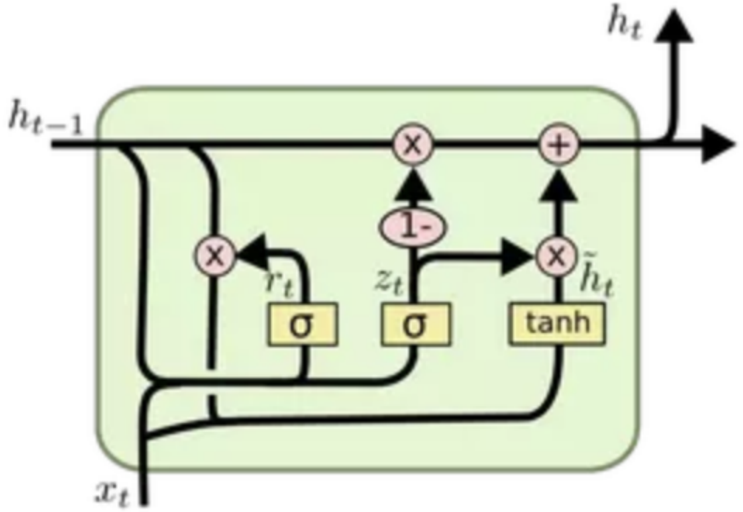
LSTM



$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right)$$
$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right)$$
$$o_t = \sigma\left(x_t U^o + h_{t-1} W^o\right)$$
$$\tilde{C}_t = \tanh\left(x_t U^g + h_{t-1} W^g\right)$$
$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \tilde{C}_t\right)$$
$$h_t = \tanh(C_t) * o_t$$

Image By research gate.

GRU



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Image By Toward Data Science

Biderectional RNN

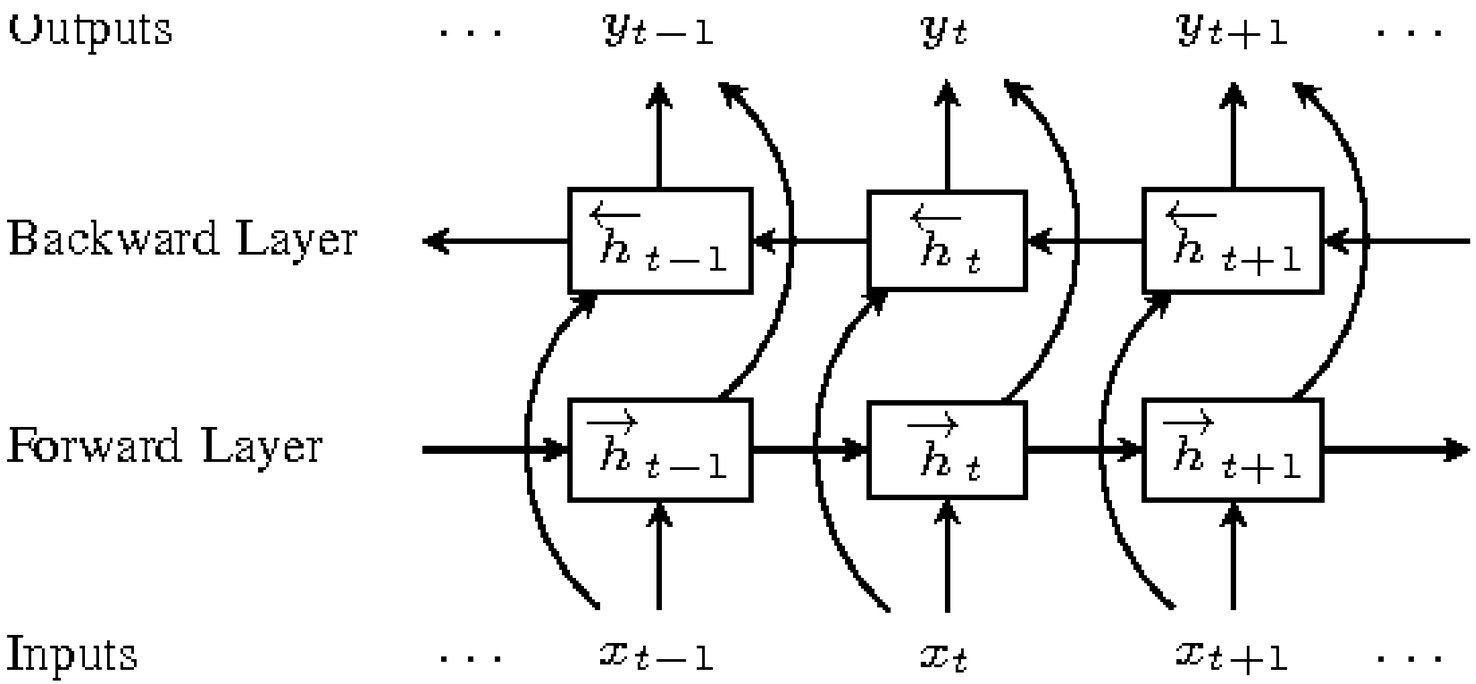


Image By https://medium.com/@madhuramiah/bi-directional-rnn-basics-of-lstm-and-gru-e114aa4779bb

Another Architecture RNN

Tree RNN

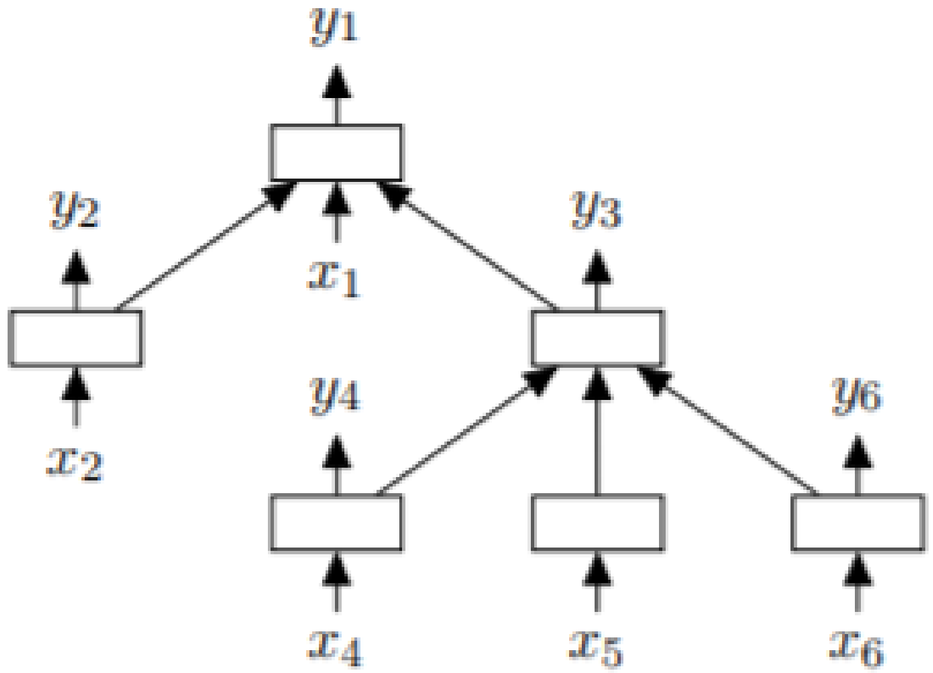


Image By Tai, Kai Sheng et al. “Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks.” ArXiv abs/1503.00075 (2015): n. pag.

Dilated RNN

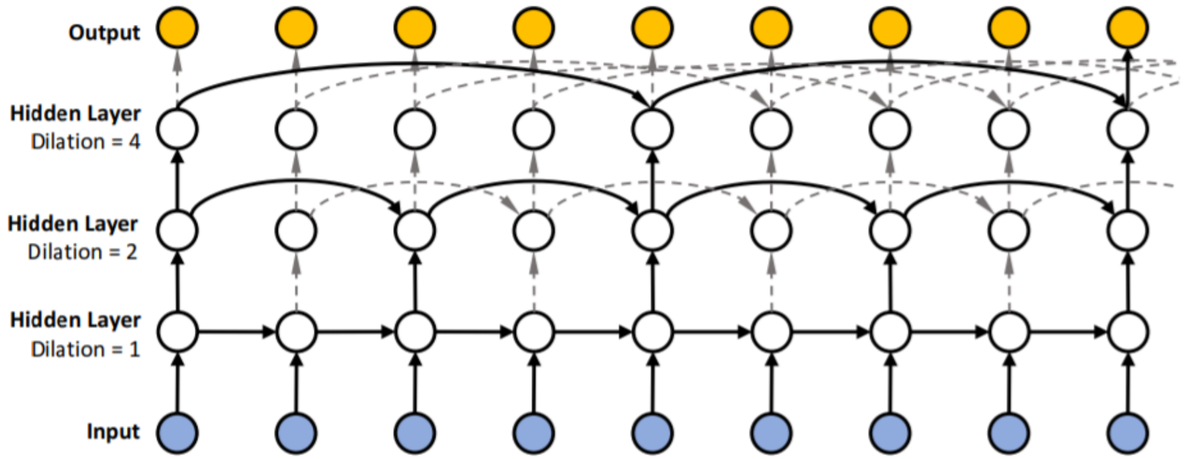


Image By Chang, S. et al. “Dilated Recurrent Neural Networks.” NIPS (2017)..

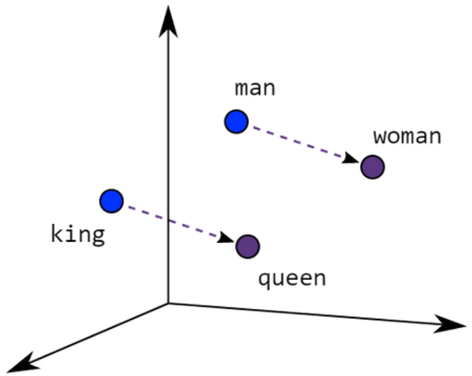
NLP

Natural Languange Processing

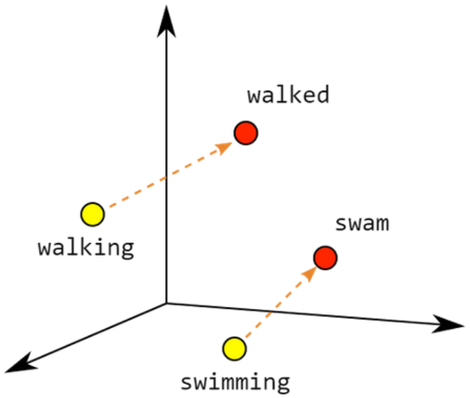
NLP Task

1. Sentiment Analysist
2. Question Answer
3. POS-tagging
4. Dialogue
5. Language Modelling
6. Machine Translation
7. DLL

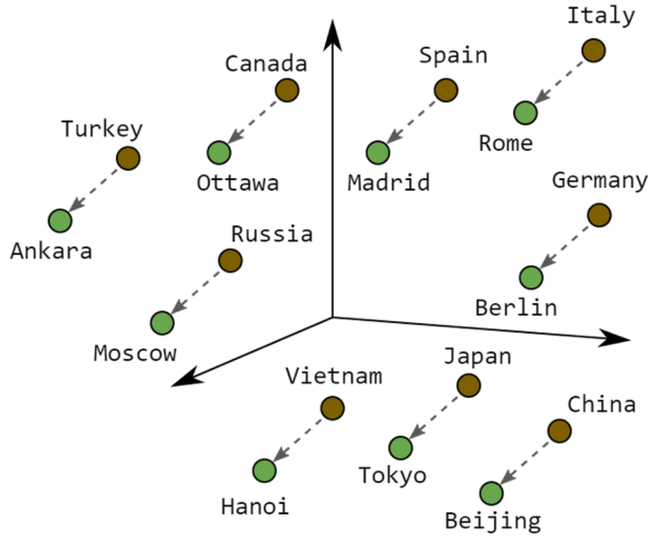
Word Embedding



Male-Female



Verb Tense



Country-Capital

Image By Google Developer

Demo

Contact

Twitter

@NaraSurya13

Email Address

ibagungnsd13@gmail.com

Telegram

@Gusagung

Next Week

Week 1 Neural Network	Week 2 CNN
Week 3 RNN	Week 4 Attention Mechanism & Transformer

See You...