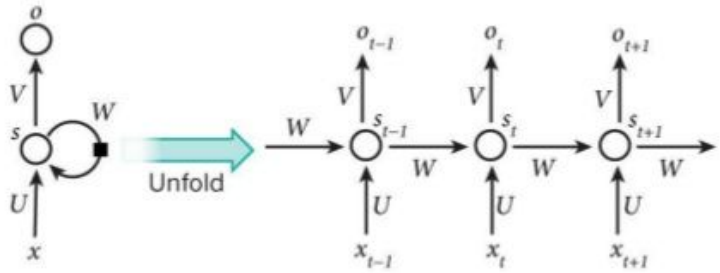
RNN's

By Mateusz Macias

Zastosowania

- Tłumaczenie
- Generowanie tekstu
- Klasyfikacje tekstu
- Automatyczne “protokołowanie” (speech2text)
- Modele języka
- Generowanie filmów
- Klasyfikacja filmów

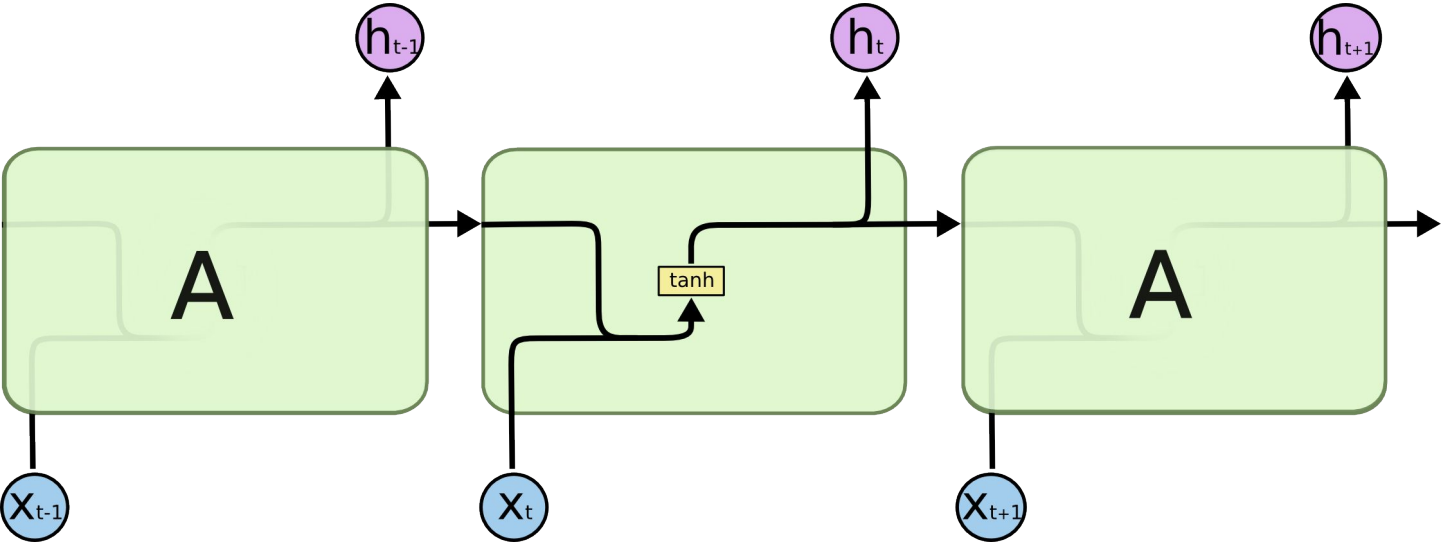
• I. Vanilla RNN



$$s_t = \tanh(Ux_t + Ws_{t-1})$$

$$\hat{y}_t = \text{softmax}(Vs_t)$$

WILDML has a series of articles to introduce RNN (4 articles, 2 GitHub repos).


h_{t-1}
h_t
h_{t+1}
A
tanh
A
X_{t-1}
X_t
X_{t+1}

Vanishing / Exploding Gradient Problem

- The issue is with the term $\frac{\partial h_t}{\partial h_k}$.
- Further maths shows (omitting many, many details):

$$\left\| \frac{\partial h_t}{\partial h_k} \right\| \leq c^{t-k}$$

- Here: c is some constant term related to θ and the choice of the activation function ϕ.
- Problem:
  - $c < 1$: Gradients tend to zero (vanish).
  - $c > 1$: Gradients will tend to infinity (explode).
- Impact of vanishing gradients to RNN: Can't "remember" impacts of long sequences.

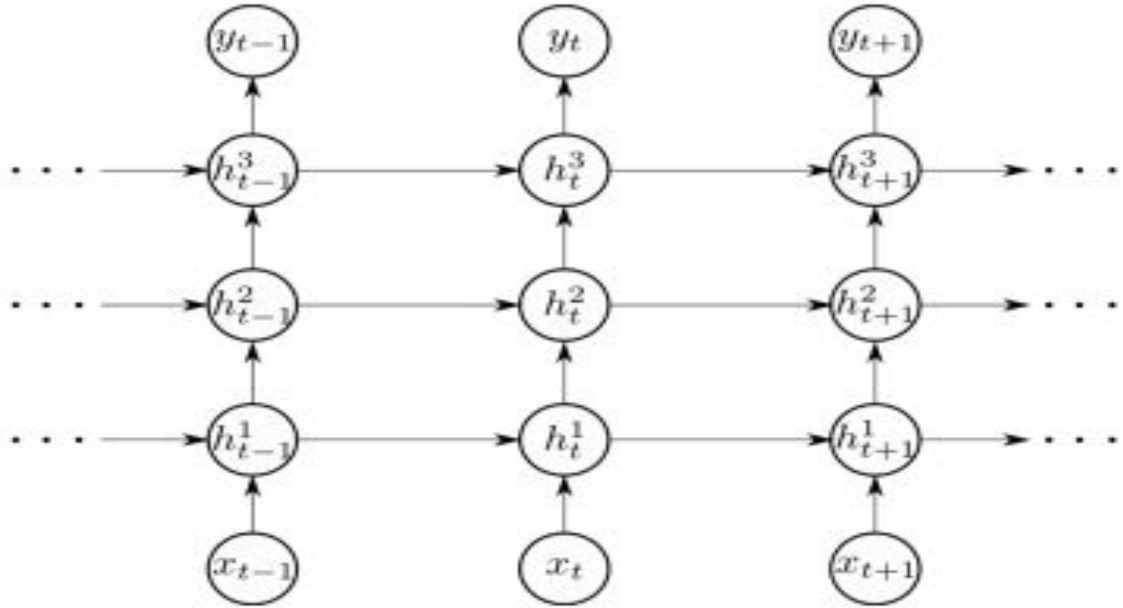
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

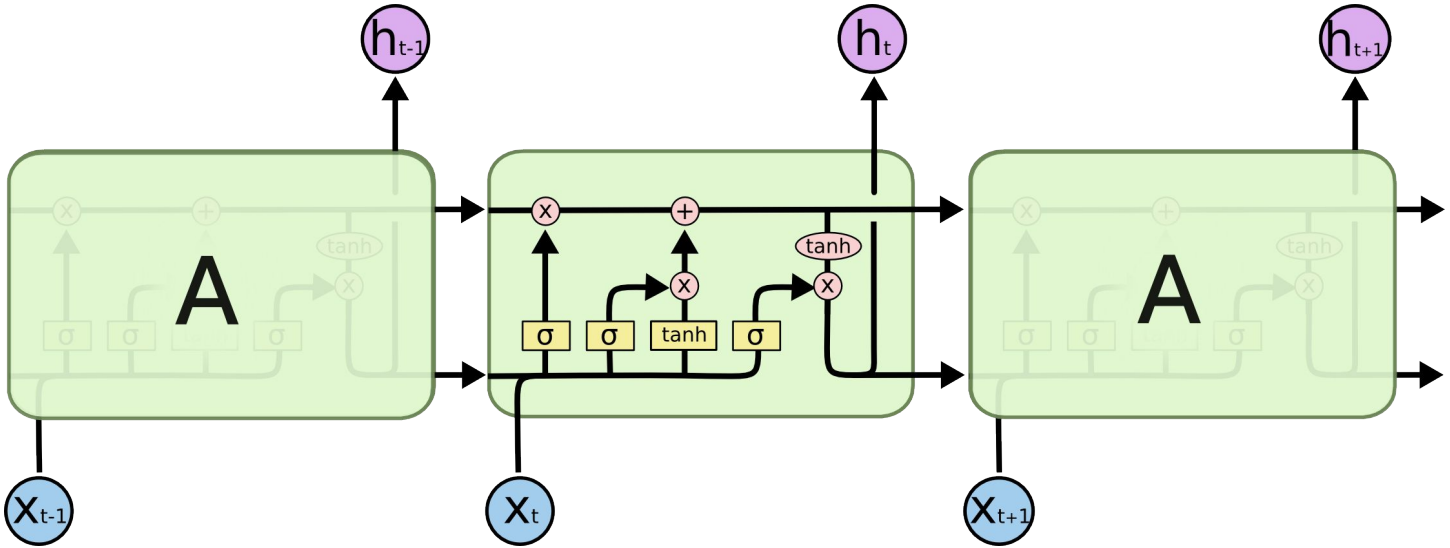
DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

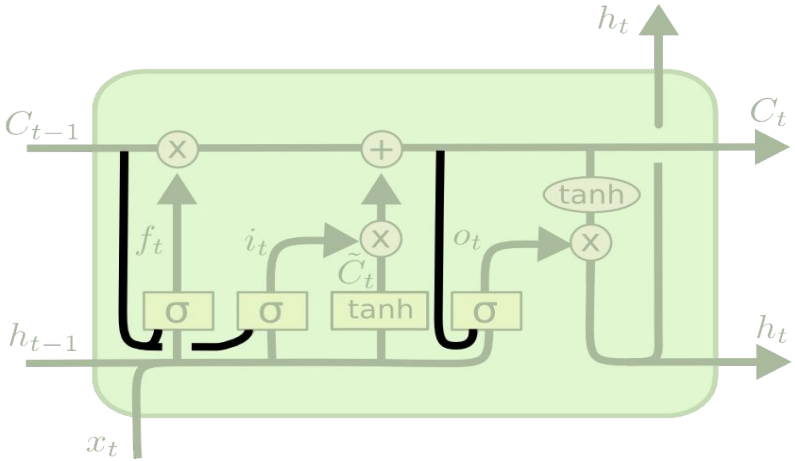
Clown:
Come, sir, I will make did behold your worship.


y_{t-1}
y_t
y_{t+1}
h^3_{t-1}
h^3_t
h^3_{t+1}
h^2_{t-1}
h^2_t
h^2_{t+1}
h^1_{t-1}
h^1_t
h^1_{t+1}
x_{t-1}
x_t
x_{t+1}

LSTM

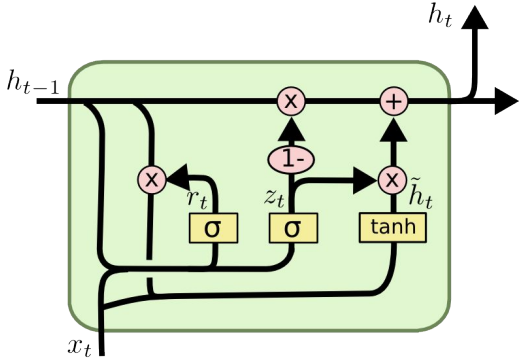


LSTM (peepholes)



$$f_t = \sigma\left(W_f \cdot [\boldsymbol{C_{t-1}}, h_{t-1}, x_t] \; + \; b_f\right)$$
$$i_t = \sigma\left(W_i \cdot [\boldsymbol{C_{t-1}}, h_{t-1}, x_t] \; + \; b_i\right)$$
$$o_t = \sigma\left(W_o \cdot [\boldsymbol{C_t}, h_{t-1}, x_t] \; + \; b_o\right)$$

GRU



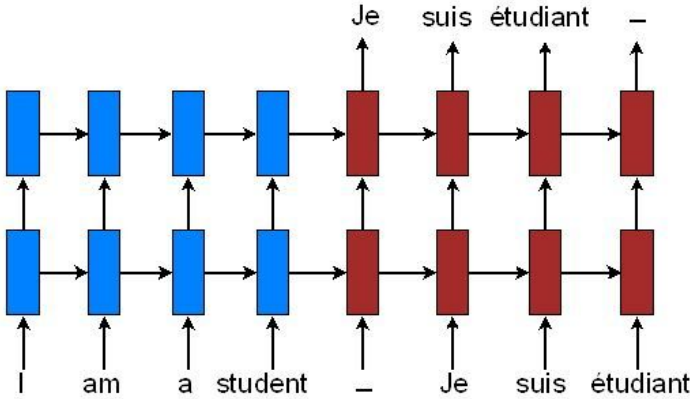
$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

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

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

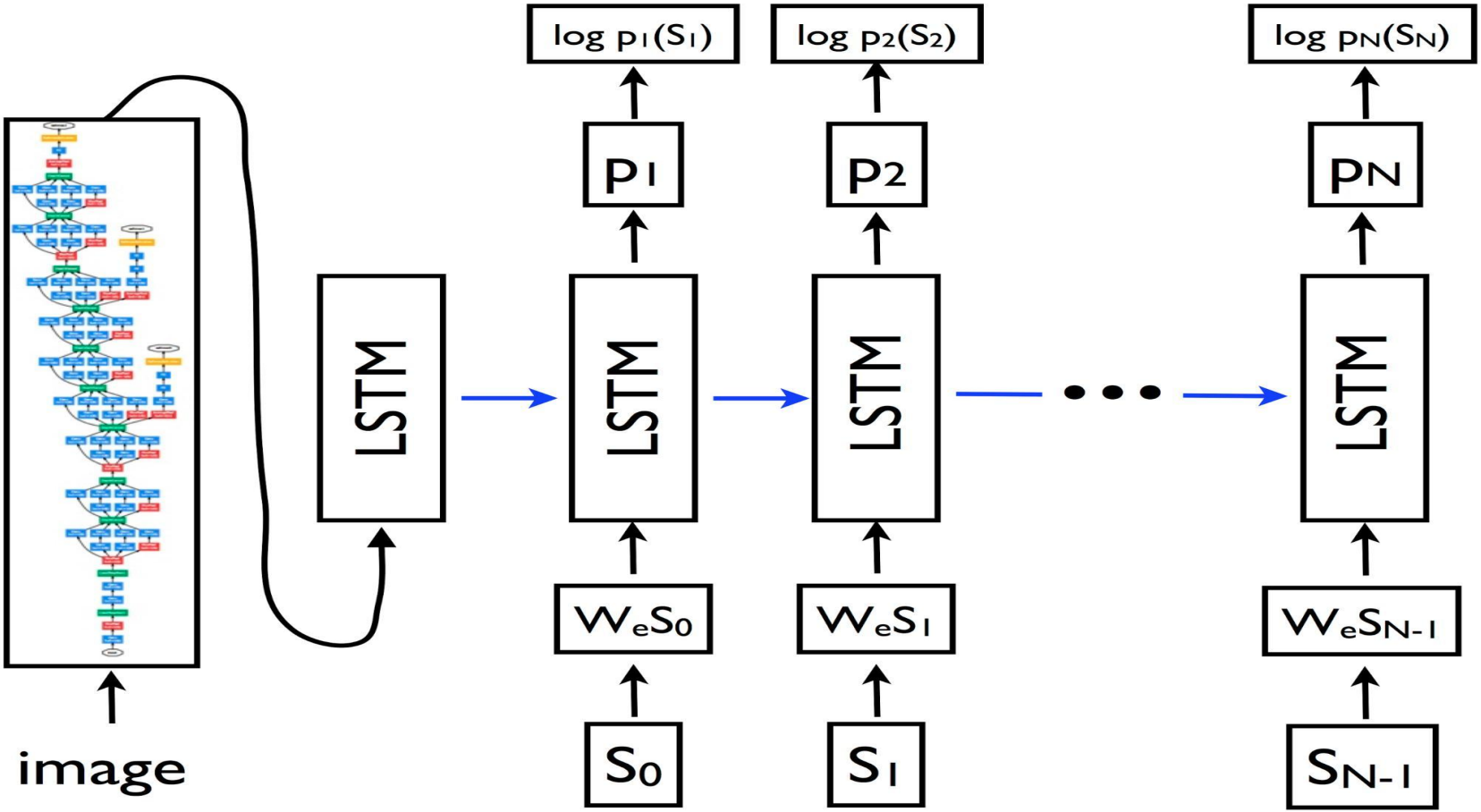
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Neural Machine Translation (NMT)



- Big RNNs trained end-to-end: encoder-decoder.
  - Generalize well to long sequences.
  - Small memory footprint.
  - Simple decoder.

Show and Tell - im2txt

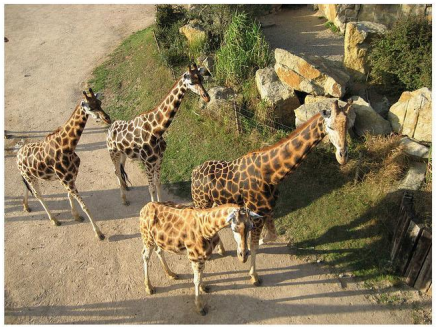


A person on a beach flying a kite.

A black and white photo of a train on a train track.

A person skiing down a snow covered slope.

A group of giraffe standing next to each other.



Źródła obrazów.

- https://image.slidesharecdn.com/rnn-lstm-161106132927/95/understanding-rnn-and-lstm-4-638.jpg?cb=1478439617
- https://qph.ec.quoracdn.net/main-qimg-b522e948f5d273af253a8a102d716680
- http://karpathy.github.io/2015/05/21/rnn-effectiveness/
- http://colah.github.io/posts/2015-08-Understanding-LSTMs/