

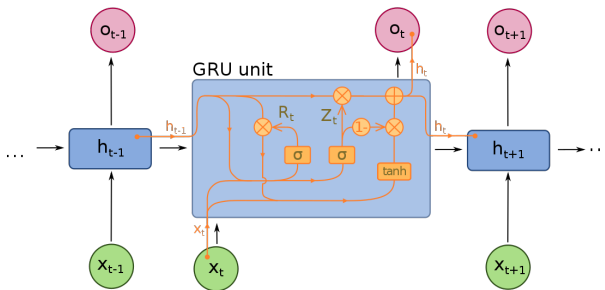
# Deep Learning par la Pratique

D'autres variantes...

---

# D'autres variantes

## Gated Recurrent Unit



$Z_t = \sigma(W_Z * x_t + U_Z * h_{t-1} + b_Z)$  (update gate  $\approx$  input et forget gates)

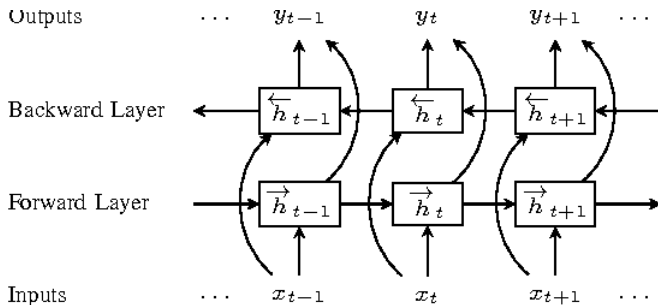
$R_t = \sigma(W_R * x_t + U_R * h_{t-1} + b_R)$  (reset gate  $\approx$  output gate)

$h_t = Z_t \circ h_{t-1} + (1 - Z_t) \circ \tanh(W_h * x_t + U_h(R_t \circ h_{t-1}) + b_h)$

Moins de paramètres que LSTM, aussi efficace dans beaucoup de tâches

## D'autres variantes

### RNN bi-directionnels



Permet de “gérer” des séquences plus longues

Dans des réseaux récurrents profonds, on peut ajouter des récurrences entre les différentes couches...

## D'autres variantes



wikiHow