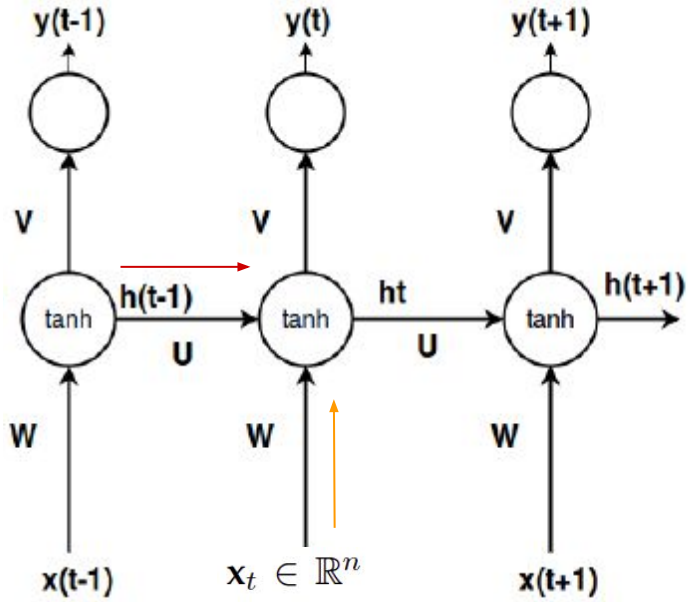
Laboratory

Recurrent Neural Network (RNN)

Deep Learning for Artificial Intelligence (DLAI)

Oscar Pina
PhD Student
Universitat Politècnica de Catalunya
oscar.pina@upc.edu

Recurrent Neural Networks



$$\mathbf{h}_t = g(\mathbf{W} \cdot \mathbf{x}_t + \mathbf{U} \cdot \mathbf{h}_{t-1} + \mathbf{b}_h)$$

LSTMs

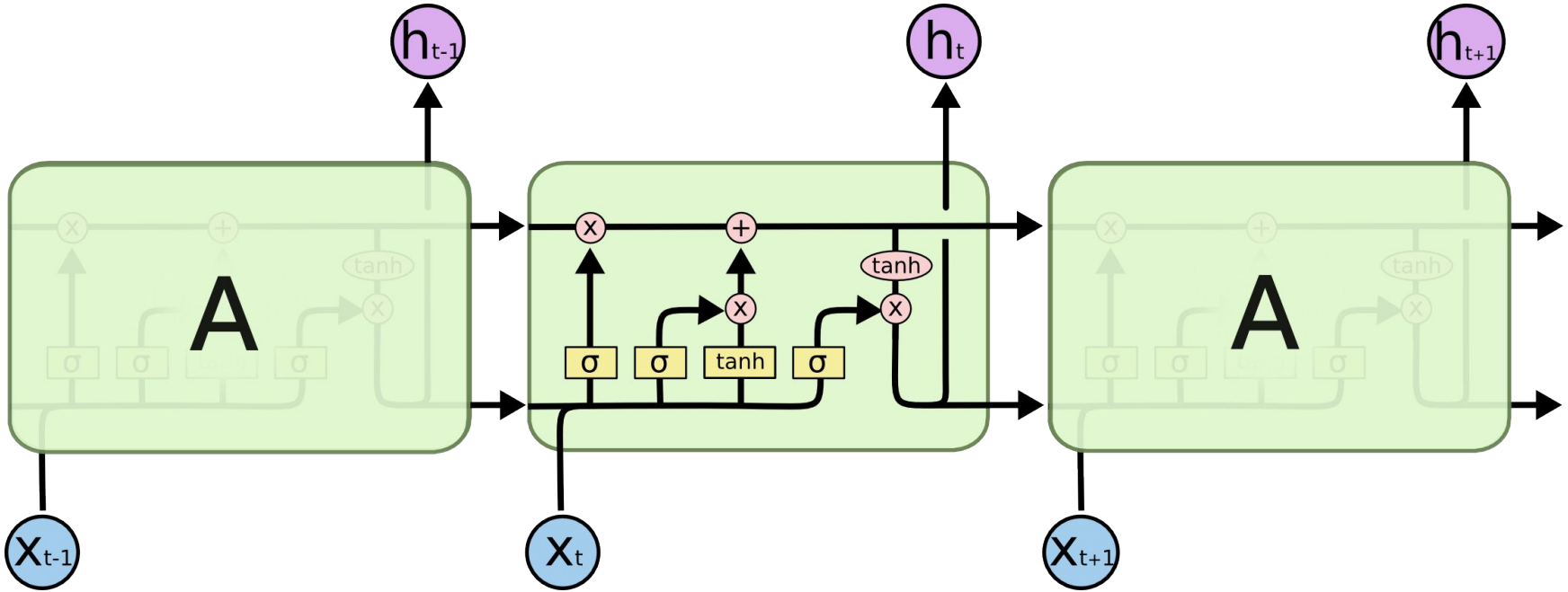


Figure: Cristopher Olah, "Understanding LSTM Networks" (2015)

LSTMs

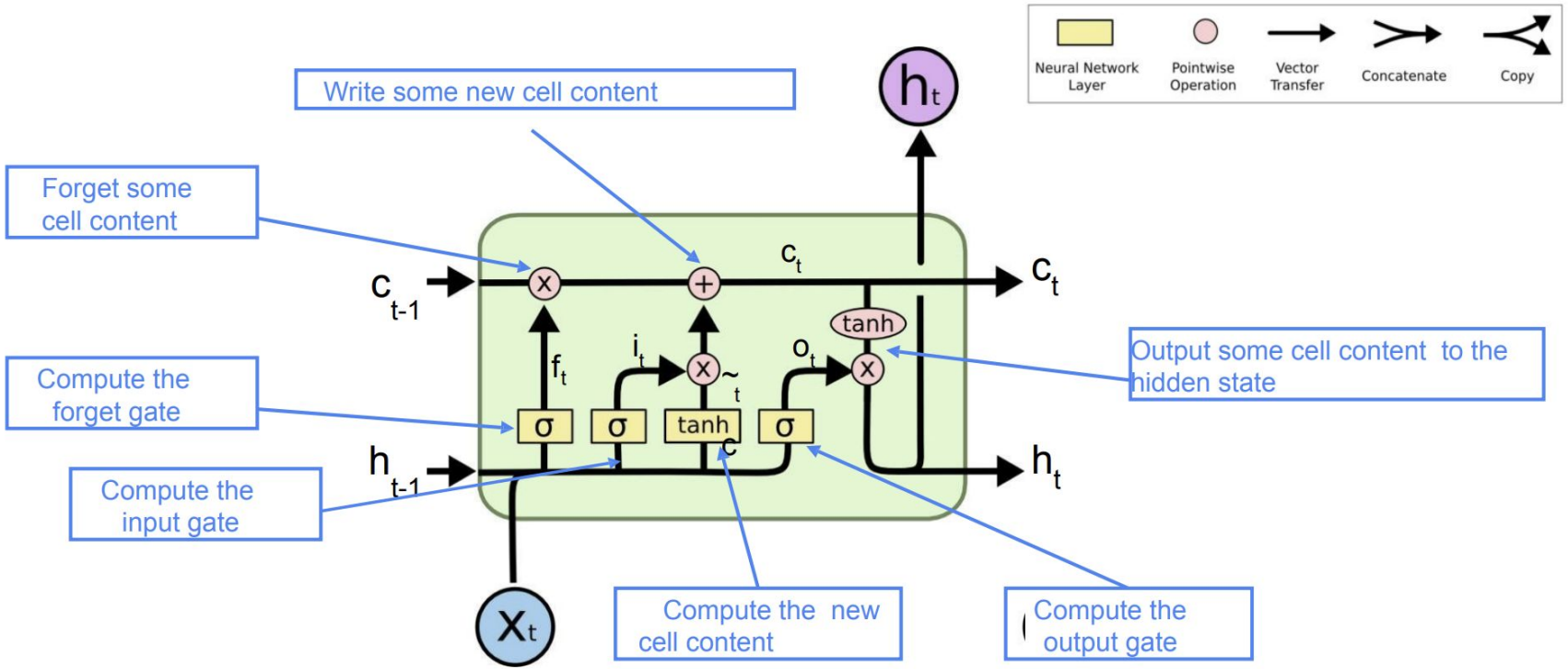
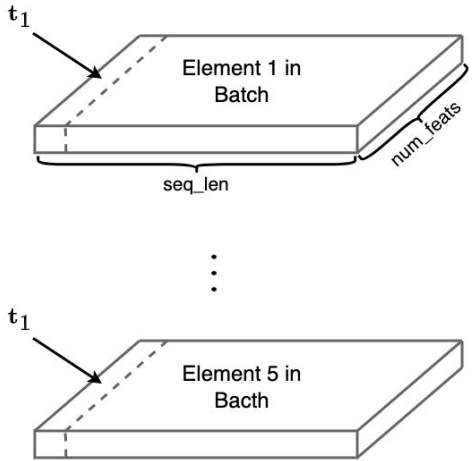


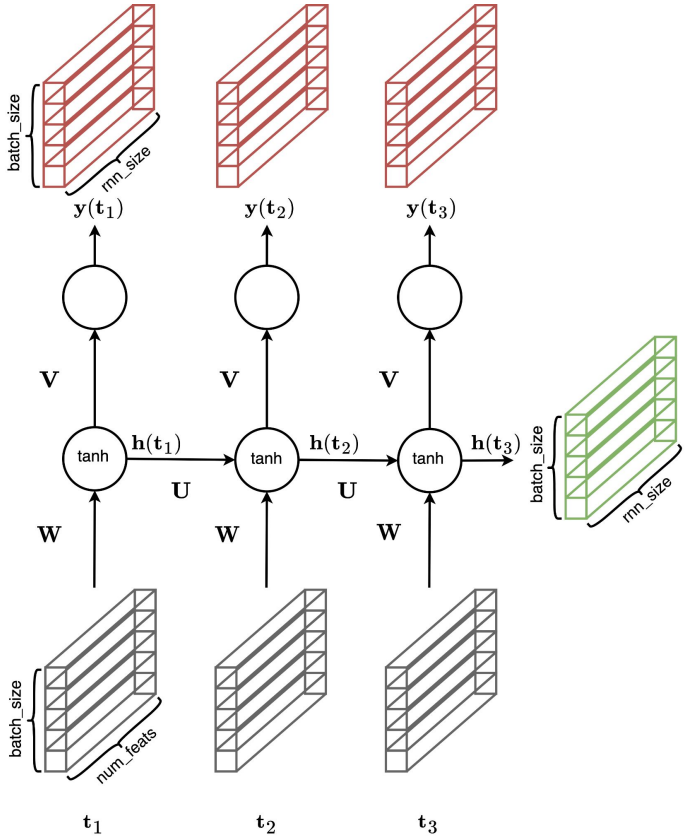
Figure: Cristopher Olah, "Understanding LSTM Networks" (2015)

Inputs of an RNN: Sequences

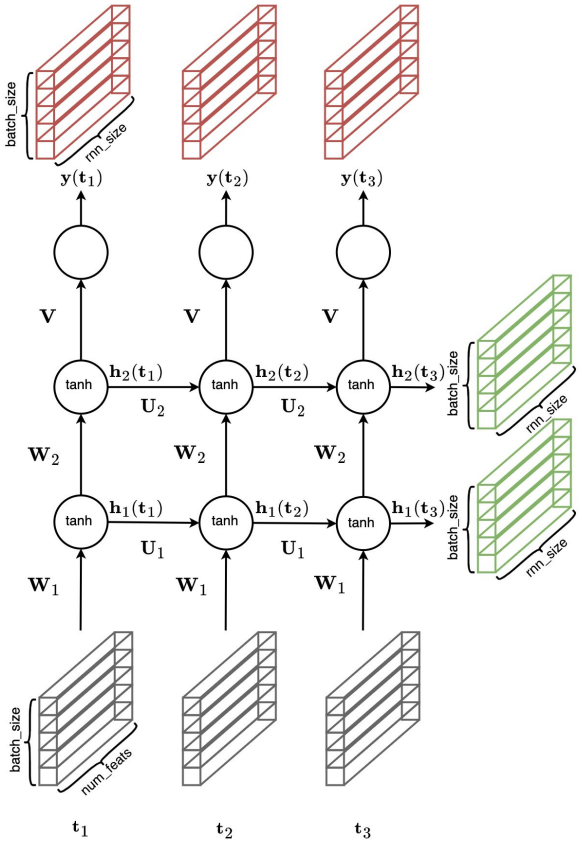
- Each input sequence [x_0, x_1, ...] has shape [seq_len, num_features].
- We will stack multiple sequences in a batch, obtaining tensors of shape [batch_size, seq_len, num_features].
- We need to use *batch_first=True* to tell PyTorch that the first dimension of the tensor is the batch dimension.



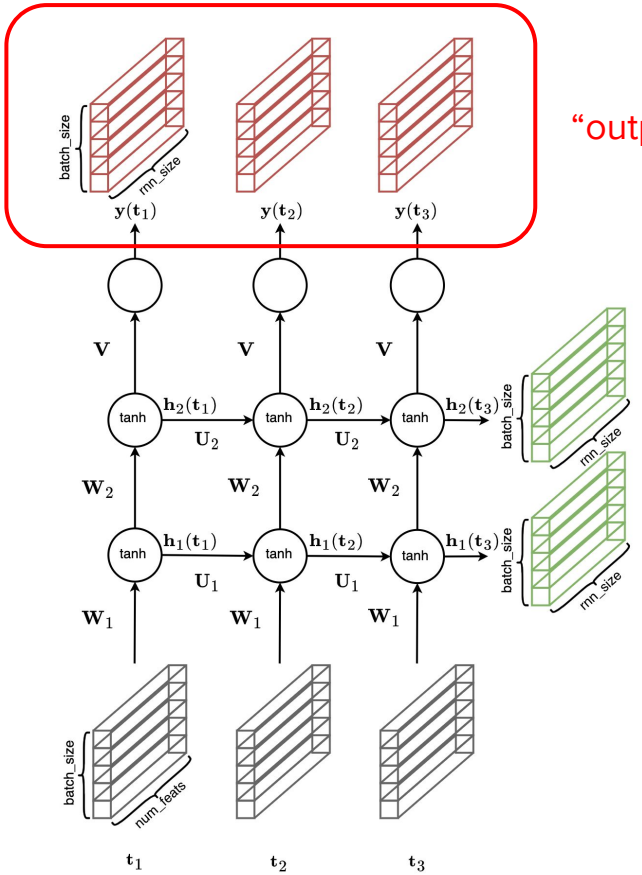
Processing inputs through RNNs (1 layer)



Processing inputs through RNNs (2 layers)



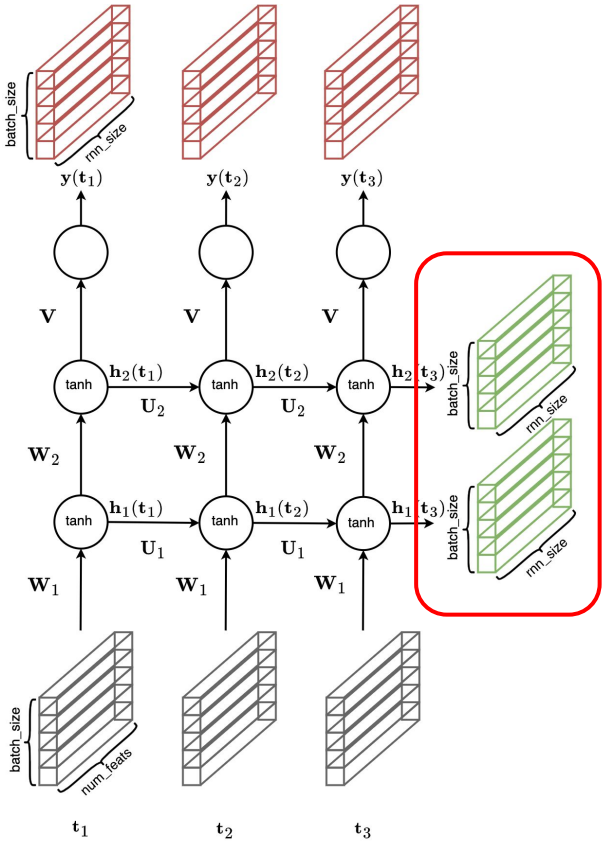
Outputs of an RNN



"output / h_t"

PYTORCH

Outputs of an RNN



"h_n / state"

PYTORCH

Task: A Character-based Language Model

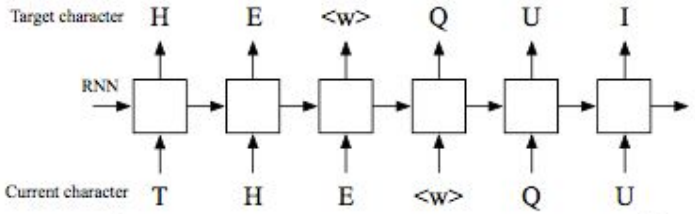


Figure: "Character Level Language Models using LSTMs", The University of British Columbia

Kick off the lab

1. Launch a web browser (Chrome recommended).
2. Login to a Google account. Create a new one if preferred.
3. Create a copy this notebook of this lab to your Google Drive. To create it, press File -> Save a copy on Drive