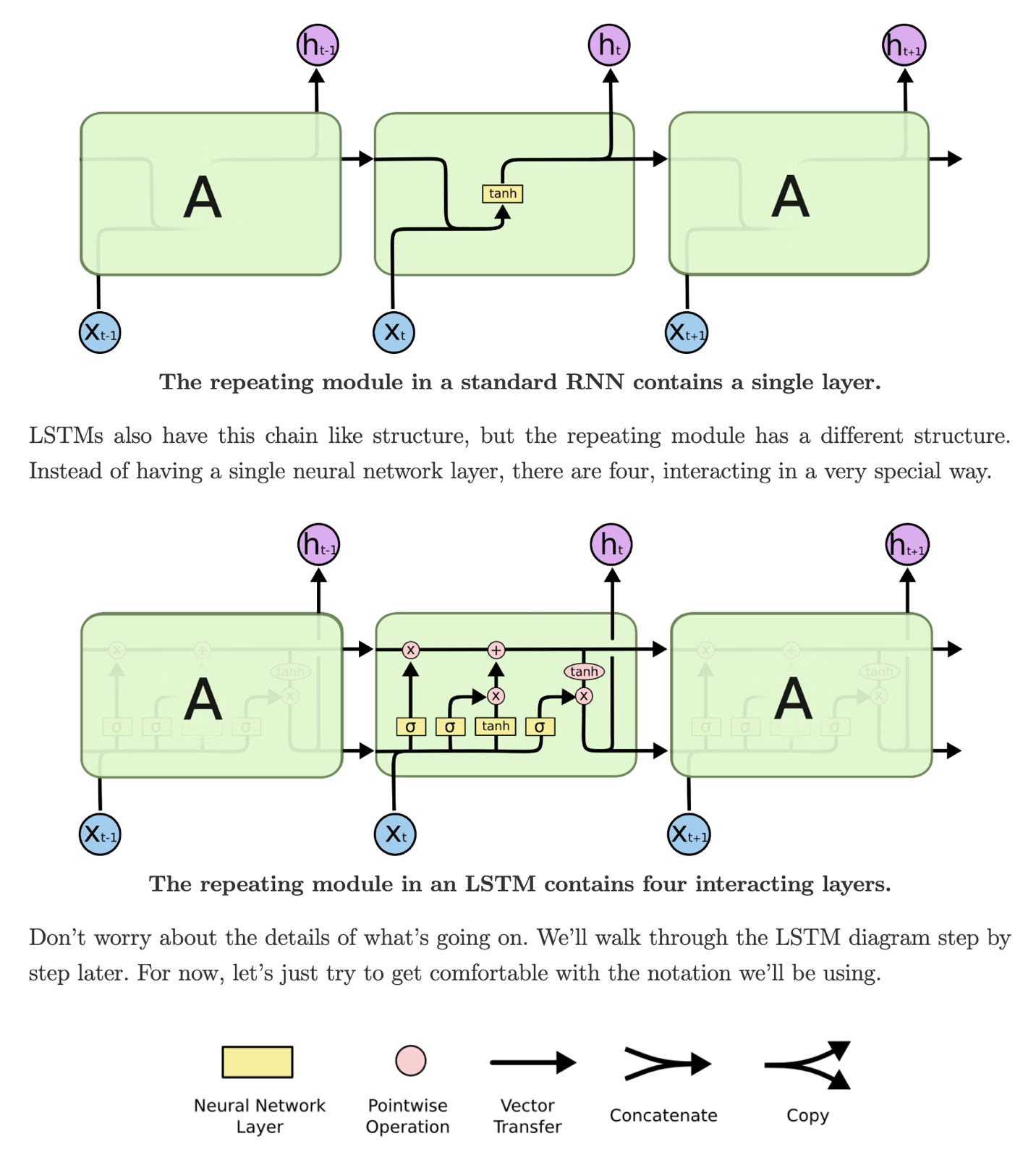
ML notes:

Fall 2019 (DeepTable project):

* **RNN/ Long-short term memory**

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>



* **Unsupervised learning: learn the distribution of examples in the feature space**

self-supervised learning: a kind of unsupervised learning, which generates labels by the model itself.

D**enoising task can be considered as a way to perform the unsupervised learning task of learning the distribution of examples in the feature space (also there are other tricks for denoising)**

* Dfencoder: (<https://towardsdatascience.com/how-to-apply-self-supervision-to-tabular-data-introducing-dfencoder-eec21c4afaef>):

An implementation of Denoising Autoenconder (DAE) for tabular data

Example 1: illustrating that trained model with swap\_p of 0.15 can extract good deep-stack features since the accuracy of predicting from wrong-labeled input is still high (0.85)

Example 2: anomaly detection

Get anomaly score for every data

Example 3: data imputation

Use anomaly score and built-in predicting methods to replace problematic data

Imputation: replace missing data