Lipton, A critical review of recurrent neural networks for sequence learning

In the setting of large datasets, simple linear models tend to under fit and often underutilize computing resources. Deep learning methods which exploit the local dependency of visual information have demonstrated record setting results on many important applications. (on CNNs)

RNNs are connectionist models with the ability to selectively pass information across sequence steps while processing sequential data one element at a time.

**Why explicitly model sequentiality?**

SVMs, logistic regression and feedforward networks have proved immensely useful without explicitly modelling time. Many models implicitly capture time by concatenating immediate predecessors and successors thus presenting a sliding window.

However not long term memory, keeping context e.g. extended dialog in a call centre automation, remembering complete context of the conversation.