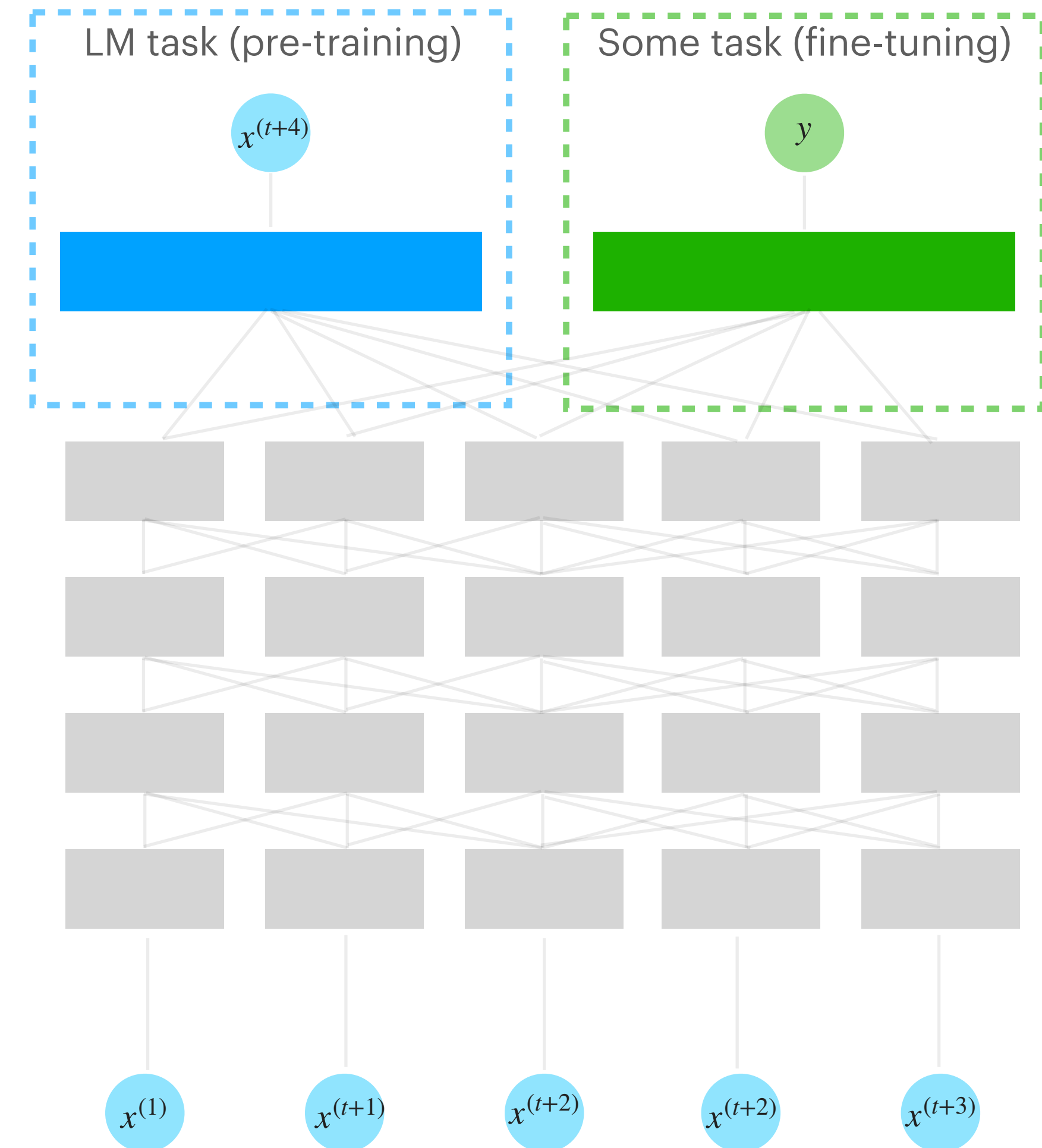# Transfer learning

- Neural language modeling provides a means to extract semantically rich features from text data. How is this then applied towards downstream tasks?
- Transfer learning:
    - LMs are first trained on next-word or next-sentence prediction tasks on a very large corpus of text (e.g., wikitext 103).
    - The output layer, which is used to map onto the word prediction output space, is removed, and replaced with a randomly initialized projection matrix mapping to the output space of the particular prediction task of interest. It is trained on a much lower number of examples from that task.

NLP Tasks: Text classification

- We've already covered this topic in Module 01. Let's take a look at the current SOTA.