

Lecture 18 - Contents

An overview of the main sections in this lecture.

Part 1

Efficient Architectures and Attention

Part 2

Mixture-of-Experts and Routing

Part 3

Advanced Sequence Models

Hands-on

Efficient Training Lab

This outline is for guidance. Navigate the slides with the left/right arrow keys.

