

Lecture 12 - Contents

An overview of the main sections in this lecture.

Part 1

Model Compression

Part 2

Distillation, Quantization, and
Pruning

Part 3

Edge and Mobile Deployment

Hands-on

Compression Hands-on

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

