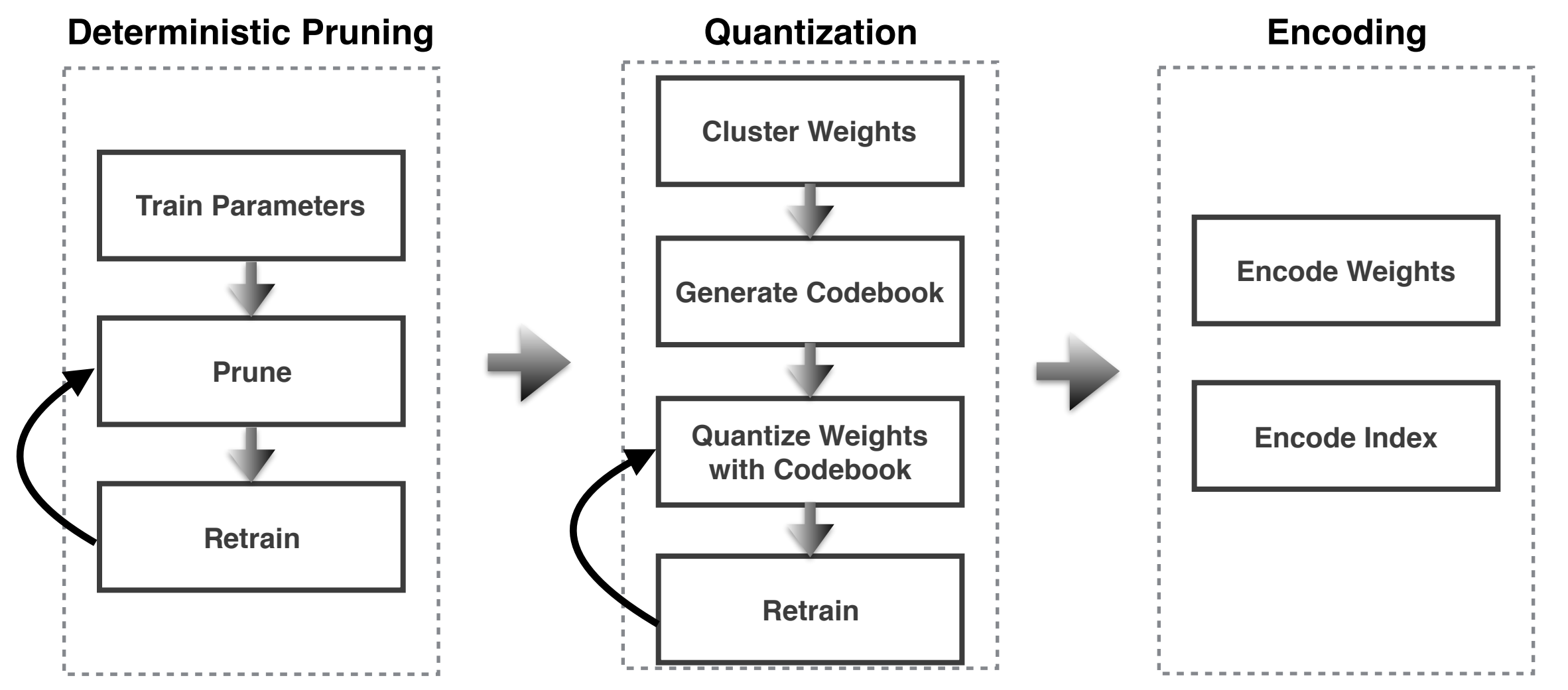
Deterministic Pruning
Train Parameters
Prune
Retrain
Quantization
Cluster Weights
Generate Codebook
Quantize Weights with Codebook
Retrain
Encoding
Encode Weights
Encode Index