

Early vs Late Training: Which Model Learns Faster?

