Training Loss over Epochs Train Loss 10^{0} 10^{-1} Loss 10^{-2} 10^{-3} 15000 20000 25000 5000 10000 30000 35000 **Epoch**