

Target: h_0 h_1 h_2 h_3 h_4 h_5 h_6 h_7 (hidden states from 1-step prediction)

Output: h'_0 h'_1 h'_2 h'_3 h'_4 h'_5 h'_6 h'_7 (hidden states from multi-step prediction)

Simplified Professor Forcing Loss