

Model	Hidden Dim.	Blocks	Params	Optim.	Batch Size	Act.	Learning Rate	Epoch
MLP	[60, 500]	[2, 8]	[11,407, 1,006,607]	Adam	4096	Swish	$3e^{-4} \times 0.7^{\lfloor e/10 \rfloor}$	50
LSTM	[25, 190]	[2, 8]	[9,757, 1,028,827]	Adam	4096	Swish	$3e^{-4} \times 0.7^{\lfloor e/10 \rfloor}$	50
GRU	[28, 220]	[2, 8]	[9,415, 1,030,507]	Adam	4096	Swish	$3e^{-4} \times 0.7^{\lfloor e/10 \rfloor}$	50
CNN	[22, 190]	[2, 8]	[9,423, 1,013,882]	Adam	4096	Swish	$3e^{-4} \times 0.7^{\lfloor e/10 \rfloor}$	50
Ours	[25, 230]	[1, 2]	[13,695, 994,247]	Adam	4096	Swish	$3e^{-4} \times 0.7^{\lfloor e/10 \rfloor}$	50