

Models	Validation NLL (V)	Test NLL (T)	Generalization (V-T)
LSTM	102.37	101.39	0.98
LSTM+	102.91	101.96	0.95
PRU	101.53	100.22	1.31
PRU+	101.06	99.79	1.27