GRU vs LSTM Performance Comparison

Metric	GRU Value	LSTM Value	Description
Initial Loss	6.3152	5.7486	Loss after the first epoch, indicating model initialization Performance.
Final Loss	0.3046	0.2662	Loss after 40 epochs; lower is better and shows Convergence.
Initial Accuracy	19.81%	19.94%	Accuracy at the first epoch; nearly identical performance.
Final Accuracy	89.98%	91.04%	Accuracy after 40 epochs; LSTM slightly out performs GRU.
Initial Val Loss	3.5069	3.5204	Validation loss in epoch 1; GRU starts off slightly better.
Final Val Loss	0.3642	0.4050	Validation loss at the last epoch; GRU generalizes Slightly better.
Initial Val Accuracy	16.67%	18.67%	Initial validation accuracy; LSTM shows a better start.
Final Val Accuracy	85.67%	86.00%	Final validation accuracy; LSTM shows a marginal advantage.
Convergence Epoch	27	31	GRU converges faster than LSTM.
Training Time	Faster	Slower	GRU requires fewer computations, making it faster.
Memory Usage	Lower	Higher	GRU is more memory efficient due to fewer gates.
Complexity	Lower	Higher	LSTM has more gates and states to compute.
Stability	Good	Excellent	LSTM is more stable in learning complex dependencies.
Generalization	Slightly Better	Good	GRU avoids overfitting slightly better as per val loss.
Prediction Quality	Coherent	Richer	GRU predictions are stable; LSTM's are more expressive.
Training Curve	Smoother	Fluctuates	GRU shows more consistent training trend.