

## What is Teacher Forcing?

### Definition

**Teacher Forcing** is a training strategy where the **ground-truth (target) tokens** from the training data are used as inputs to the decoder at each time step, instead of the model's own predictions.

#### Without Teacher Forcing

(Autoregressive)

Step 1:



#### With Teacher Forcing

(Training Mode)

Step 1:





↳ Uses own predictions (errors accumulate!)

- ✗ Errors compound over time
- ✗ Slower training convergence
- ✗ Exposure bias problem



↳ Uses ground-truth targets (no error accumulation!)

- ✓ Faster training convergence
- ✓ Prevents error accumulation
- ✓ More stable gradient flow



Teacher Forcing is used during **training** to speed up learning. During **inference**, the model uses its own predictions (autoregressive).