The Transformer Pipeline: Input → Process → Output

INPUT
"The cat sat"

MODEL
(Transformer)

OUTPUT
"Le chat..."

1. Words enter
INPUT:
7 words
(text)

2. Compute
3. Predict
PROCESS:
All words
at once!

OUTPUT:
7 words
(translation)

WHY: We need to understand what each word means in context to produce correct output

Pipeline Flow