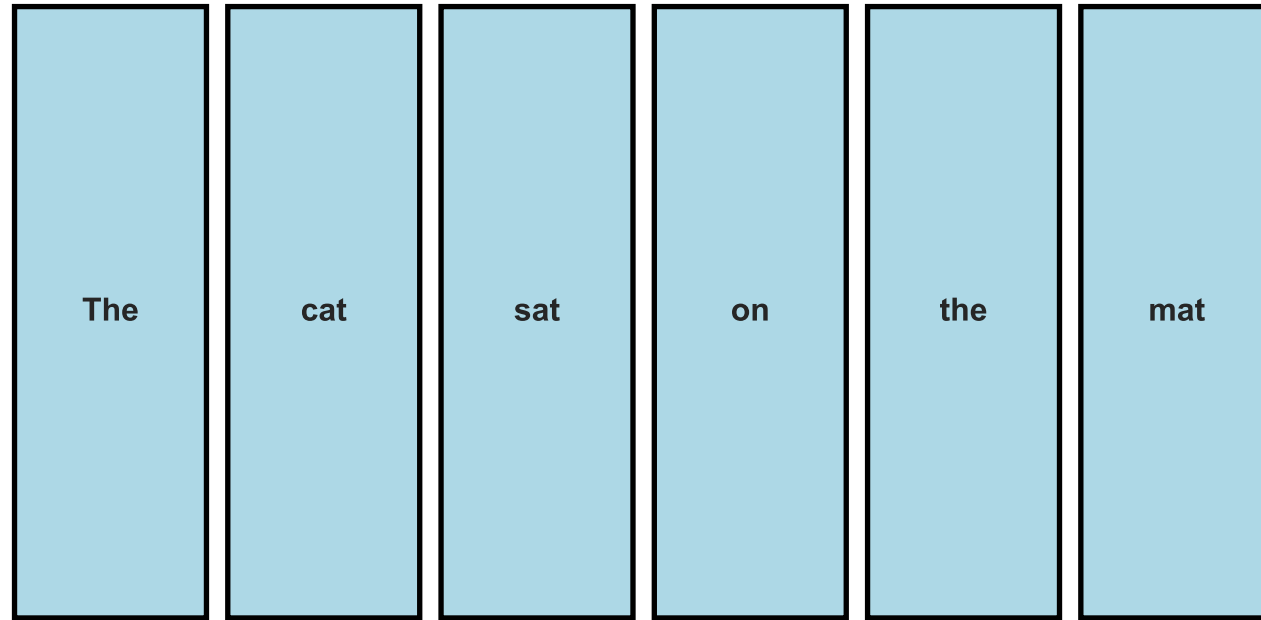
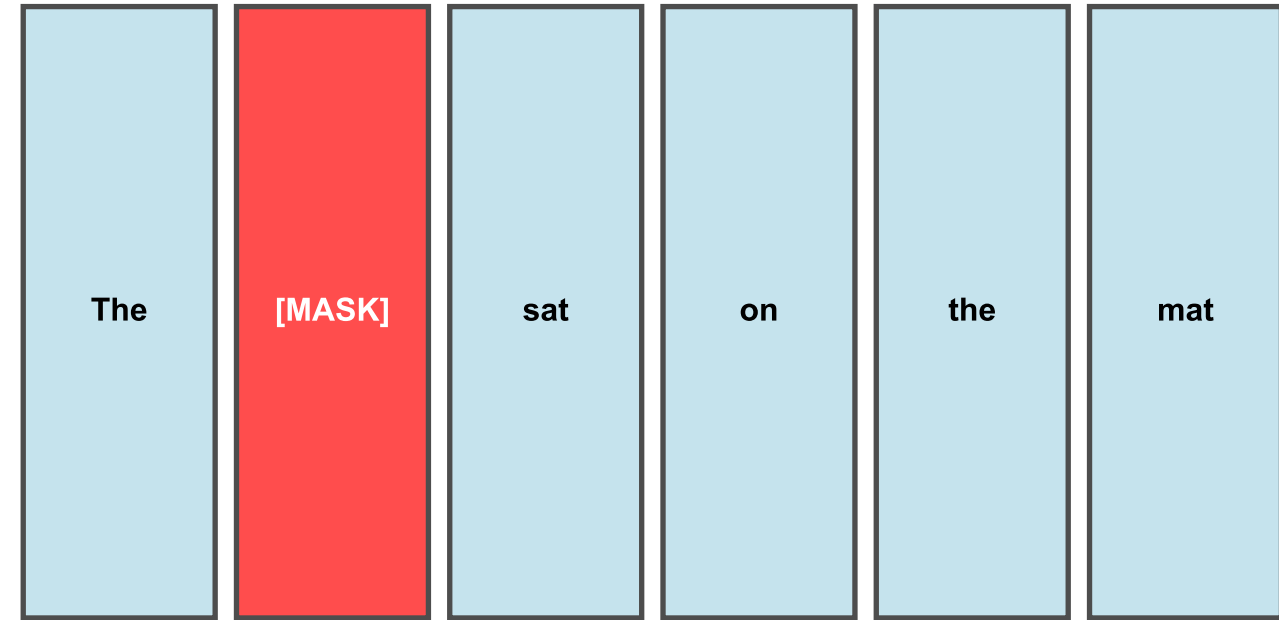


BERT Masked Language Modeling Process

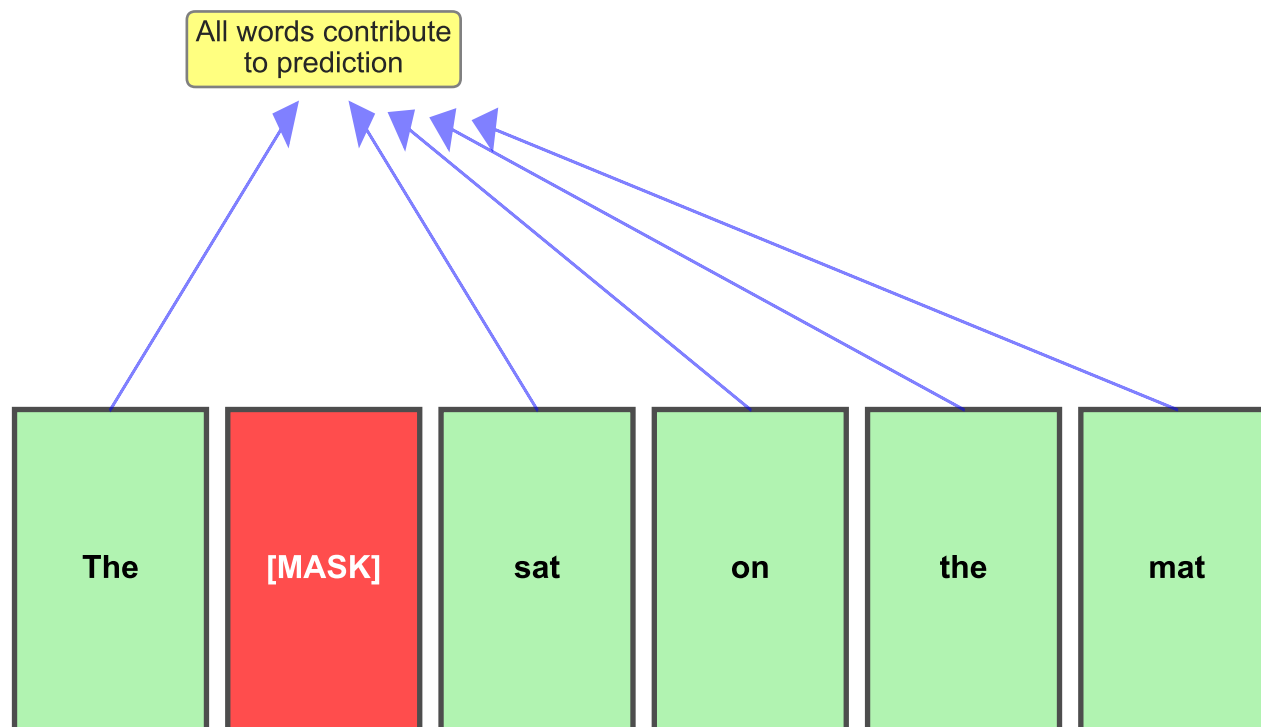
Step 1: Original Sentence



Step 2: Apply Masking (15% randomly)



Step 3: Bidirectional Context



Step 4: Predict Masked Token

