

The “The cat is on the” cat is on the

Input

Initial
Embeddings

Contextual
Embeddings

Layer 1

Layer 2

Output

Next Token Probabilities

