Mixture Distribution

Output Embedding: 1′ 2′ 3′ 4′ 5′ (left); 2′ 3′ 4′ 5′ 6′ (right)

Encoder-only Transformer

Decoder-only Transformer

Patch Embedding: 1 2 3 4 mask (left); 1 2 3 4 5 (right)

Input Time Series