

```

EncoderDecoder(
  (encoder): Encoder(
    (layers): ModuleList(
      (0): EncoderLayer(
        (self_attn): MultiHeadedAttention(
          (linears): ModuleList(
            (0): Linear(in_features=512, out_features=512, bias=True)
            (1): Linear(in_features=512, out_features=512, bias=True)
            (2): Linear(in_features=512, out_features=512, bias=True)
            (3): Linear(in_features=512, out_features=512, bias=True)
          )
          (dropout): Dropout(p=0.1, inplace=False)
        )
        (feed_forward): PositionwiseFeedForward(
          (w_1): Linear(in_features=512, out_features=2048, bias=True)
          (w_2): Linear(in_features=2048, out_features=512, bias=True)
          (dropout): Dropout(p=0.1, inplace=False)
        )
        (sublayer): ModuleList(
          (0): SublayerConnection(
            (norm): LayerNorm()
            (dropout): Dropout(p=0.1, inplace=False)
          )
          (1): SublayerConnection(
            (norm): LayerNorm()
            (dropout): Dropout(p=0.1, inplace=False)
          )
        )
      )
      (1)...(5): total of 6 'EncoderLayer' layers, same as above
    )
    (norm): LayerNorm()
  )
  (decoder): Decoder(
    (layers): ModuleList(
      (0): DecoderLayer(
        (self_attn): MultiHeadedAttention(
          (linears): ModuleList(
            (0): Linear(in_features=512, out_features=512, bias=True)
            (1): Linear(in_features=512, out_features=512, bias=True)
            (2): Linear(in_features=512, out_features=512, bias=True)
            (3): Linear(in_features=512, out_features=512, bias=True)
          )
          (dropout): Dropout(p=0.1, inplace=False)
        )
        (src_attn): MultiHeadedAttention(
          (linears): ModuleList(
            (0): Linear(in_features=512, out_features=512, bias=True)
            (1): Linear(in_features=512, out_features=512, bias=True)
            (2): Linear(in_features=512, out_features=512, bias=True)

```

```

        (3): Linear(in_features=512, out_features=512, bias=True)
    )
    (dropout): Dropout(p=0.1, inplace=False)
)
(feed_forward): PositionwiseFeedForward(
    (w_1): Linear(in_features=512, out_features=2048, bias=True)
    (w_2): Linear(in_features=2048, out_features=512, bias=True)
    (dropout): Dropout(p=0.1, inplace=False)
)
(sublayer): ModuleList(
  (0): SublayerConnection(
    (norm): LayerNorm()
    (dropout): Dropout(p=0.1, inplace=False)
  )
  (1): SublayerConnection(
    (norm): LayerNorm()
    (dropout): Dropout(p=0.1, inplace=False)
  )
  (2): SublayerConnection(
    (norm): LayerNorm()
    (dropout): Dropout(p=0.1, inplace=False)
  )
)
)
)
(1)...(5): total of 6 'DecoderLayer' layers, same as above
)
(norm): LayerNorm()
)
(src_embed): Sequential(
  (0): Embeddings(
    (lut): Embedding(11, 512)
  )
  (1): PositionalEncoding(
    (dropout): Dropout(p=0.1, inplace=False)
  )
)
)
(tgt_embed): Sequential(
  (0): Embeddings(
    (lut): Embedding(11, 512)
  )
  (1): PositionalEncoding(
    (dropout): Dropout(p=0.1, inplace=False)
  )
)
)
(generator): Generator(
  (proj): Linear(in_features=512, out_features=11, bias=True)
)
)
)

```