

Softmax

DECODE

Linear

c

Attention Weights

$$\alpha = (\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n)$$

+

$\alpha_1 \cdot h_1$

$\alpha_2 \cdot h_2$

$\alpha_3 \cdot h_3$

\dots

$\alpha_n \cdot h_n$

h_1

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

Bidirectional LSTM

Enc. Output
(h_1)

ENCODE

Linear

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

Bidirectional LSTM

Embedding Layer

w
o
r
d
1

w
o
r
d
2

w
o
r
d
3

w
o
r
d
...

w
o
r
d
t