

6: Target label

5: Multiple blocks of self-attention and position-wise FFN layers.

4: Inner attention

3: Multiple blocks of LSTM and high-way layers.

2: bi-LSTM

1: Sequence of MFCC,Energy, Pitch

