

Tasks	Section	Class	criterion	Max mark	Sub-Total	File
1	Building Blocks	LayerNorm	:forward	2	12	
2		MultiHeadAttention	:attention	4		
3			:forward	5		
4		FeedForwardLayer	:forward	1		
5	Architecture	TransformerEncoderLayer	:pre_layer_norm_forward	2	20	<i>a2_transformer_model.py</i>
6			:post_layer_norm_forward	1		
7		TransformerDecoderLayer	:__init__	4		
8			:pre_layer_norm_forward	2		
9		:post_layer_norm_forward	2			
10		TransformerDecoder	:forward	3		
11		TransformerEncoderDecoder	:create_pad_mask	1		
12			:create_causal_mask	2		
13	:forward		3			
15	Decoding: Greedy, and beam-search	TransformerEncoderDecoder	:greedy_decode	5	20	
17			: expand_encoder_for_beam_search	3		
18			: repeat_and_reshape_for_beam_search	1		
19			: initialize_beams_for_beam_search	6		
20			: pad_and_score_sequence_for_beam_search	3		
21			:finalize_beams_for_beam_search	2		
22	Training and testing	TransformerRunner	:train_input_target_split	1	12	<i>a2_transformer_runner.py</i>
23			:train_step_optimizer_and_scheduler	1		
24			:train_for_epoch	5		
25			:translate	2		
26			:compute_batch_total_bleu	3		
27	BLEU score		BLEU score: grouper	2	8	<i>a2_bleu_score.py</i>
28			BLEU score: n_gram_precision	2		
29			BLEU score: brevity_penalty	2		
30			BLEU score: BLEU_Score(...)	2		
32	Analysis			8	8	<i>analysis.pdf</i>
			<b>Total</b>	80	80	