Predicted errors	V	X	X	V	V	X	V	$\bigwedge$
Error detection layer	Tagging Layer							
Token rep	E0	E1	E2	E3	E4	E5	E6	
Encoder	BERT							
Confidence embeds	CE-S	CE1	CE2	CE3	CE4	CE5	CE-S	
Position embeds	+	+	+	+	+	+	+	
Segment embeds	+	+	+	+	+	+	+	
Token embeds	Т	<b>T</b>	<b>T</b>	<b>T</b>	<b>T</b>	<b>T</b>	<b>T</b>	
Input tokens	[CLS]	t1	t2	t3	t4	t5	[SEP]	
Confidence scores	[CLS]	<b>c1</b>	<b>c2</b>	с3	c4	<b>c5</b>	[SEP]	