	Casidas come solvo as of	Bode	h	Non	naliz	adion	
	Consider came cetyp as of Bodch Normalization.						
	(Z ₁)	t'	fz	2,	72	Zz	
	51	2	3	7	5	4	> H'E'
	(22) 12 9	1	1	2	3	4	2 H2 52
	f. 2.	5	4	1	2.	3	>11.62
	Z ₃	6	1	7	5	6	> My 64
	Y3B3	7)	3	3	4	34.6
							7.5.5
1	To be accordination to a one colculated across features						
1	In layer normalization M & o are calculated across features (row wise) & are readjusted using the Y & B values of the respective columns (values of the preactivation) (Z, > T, B, Z, > T,B2,						
1	(row wise) A we readposted by the proposition (7 -> TB: Z-> TB:						
1	respective columns (Values or the president) (=1 1/11) 2 210)						
+	$Z_3 \rightarrow Y_3 \beta_2$						
+					(
1	In some way too mnormalizing the values of text embedding						
		110.					0

motrix we use Layer Normalization Here, as we can see, each embedding 6.5 2.41. 3.21 MIGI Probad 2:21 0.4 3.6 , 4262 vertor has it's own statistical LPad> 0 × M3 63 representation (M & 5) & the zero LPad> 0 0 My 64 0 padding embedding vectors are not How 7.5 9.2 1.5 Mg 65 affecting the true statistical 2.2 1.1. 6.7 PLG 56 are representations of the embedding 400 9 47 67 2.9 madrix. today 23 6.5 00

Eq. for 1st dimension of embedding vector of word "How" -7.5- H5 = 0.7 -> Hypothetical & Row wise 45 & Column wise T, & B,

Readjustment: 0.77, + B,