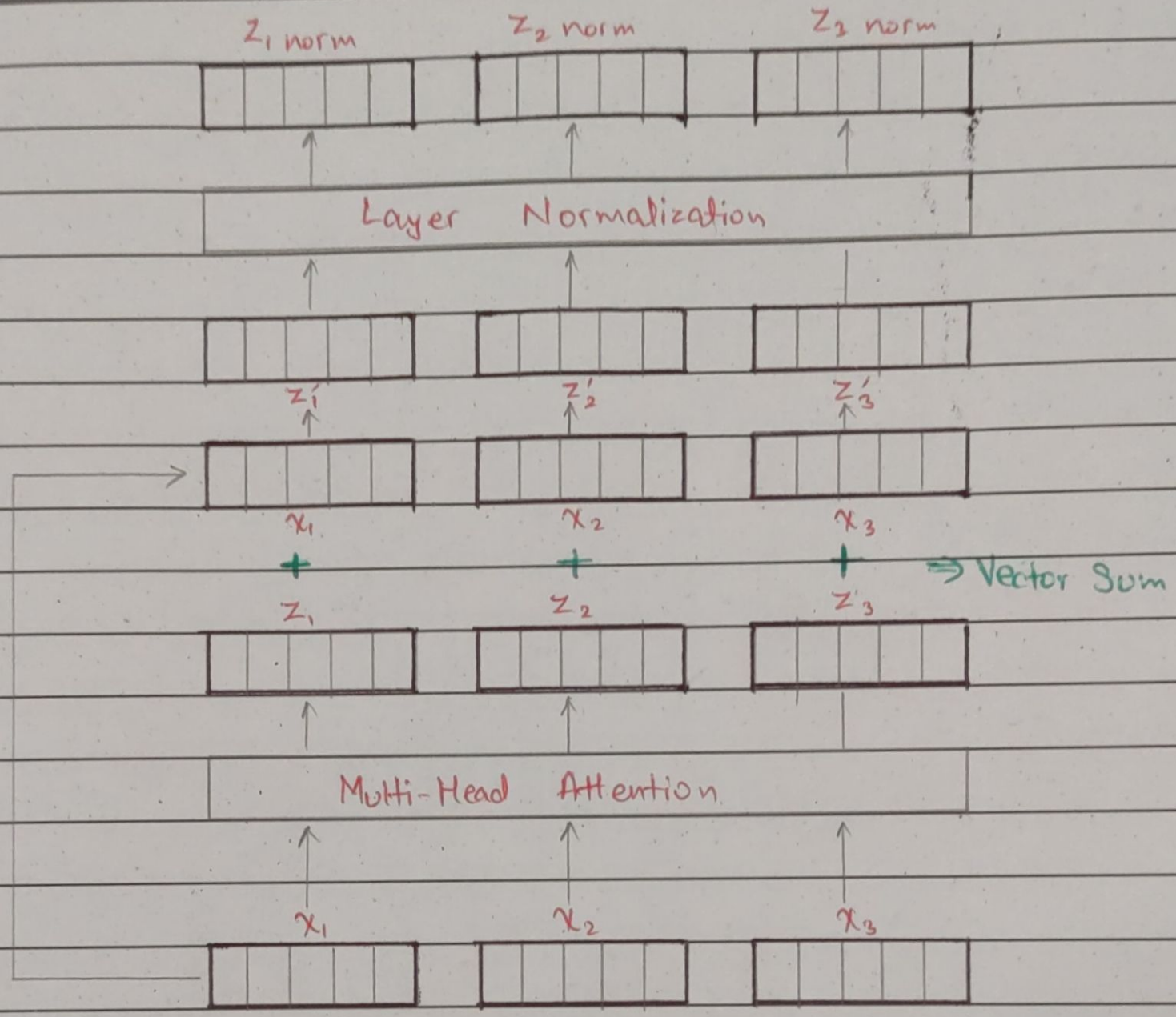

* The position encoded embedding vectors x_1, x_2 & x_3 are then passed to the Multi-Head Attention.
* The Multi-Head Attention converts the embedding vectors x_1, x_2 & x_3 into contextual embedding vectors z_1, z_2 & z_3.
* Through a residual/skip connection original x_1, x_2 & x_3 are added to z_1, z_2 & z_3 using vector addition resulting in z_1', z_2', & z_3'.
* z_1', z_2', & z_3' are then passed through Layer Normalization to normalize embedding values, resulting in $z_{1\,norm}$, $z_{2\,norm}$, & $z_{3\,norm}$.
* Dimensions of all the vectors in this stage is 512.