

Padding Attention Mask

	We	love	pie	<PAD>
We	0.2	0.1	0.7	0
love	0.2	0.6	0.2	0
pie	0.5	0.2	0.3	0
<PAD>	0	0	0	0

A

\times

$$\begin{bmatrix} - & \mathbf{V}_1 & - \\ - & \mathbf{V}_2 & - \\ - & \mathbf{V}_3 & - \\ - & \mathbf{V}_4 & - \end{bmatrix}$$

V

$=$

$$\begin{bmatrix} 0.2\mathbf{V}_1 + 0.1\mathbf{V}_2 + 0.7\mathbf{V}_3 + 0\mathbf{V}_4 \\ 0.2\mathbf{V}_1 + 0.6\mathbf{V}_2 + 0.2\mathbf{V}_3 + 0\mathbf{V}_4 \\ 0.5\mathbf{V}_1 + 0.2\mathbf{V}_2 + 0.3\mathbf{V}_3 + 0\mathbf{V}_4 \\ 0\mathbf{V}_1 + 0\mathbf{V}_2 + 0\mathbf{V}_3 + 0\mathbf{V}_4 \end{bmatrix}$$

OUTPUT