

Diagram illustrating a matrix multiplication operation:

The first matrix is a block matrix with dimensions $N \times HD_v$. It is composed of blocks $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_H$.

This matrix is multiplied (\times) by a second matrix with dimensions $HD_v \times D$, labeled $\mathbf{W}^{(o)}$.

The result is a matrix with dimensions $N \times D$, labeled \mathbf{Y} .

$$\begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \dots & \mathbf{H}_H \end{bmatrix}_{N \times HD_v} \times \mathbf{W}^{(o)}_{HD_v \times D} = \mathbf{Y}_{N \times D}$$