

$$\hat{Y} = \sigma \left(uv^{\top} \sigma \left(W_1 \sigma \left(W_2 \dots W_p \right) \right) \right) X$$