$$\hat{y}^{(1)} = \text{softmax}\left(\underset{(1,D_h)}{h^{(1)}} * \underset{(D_h,|V|)}{U} + \underset{(1,|V|)}{b_2} \right)$$

$$\hat{y}^{(2)} = \text{softmax}\left(\underset{(1,D_h)}{h^{(2)}} * \underset{(D_h,|V|)}{U} + \underset{(1,|V|)}{b_2} \right)$$

$$\hat{y}^{(3)} = \text{softmax}\left(\underset{(1,D_h)}{h^{(3)}} * \underset{(D_h,|V|)}{U} + \underset{(1,|V|)}{b_2} \right)$$

$$\underset{(1,D_h)}{h^{(0)}} \qquad h^{(1)} = \sigma\left(\underset{(1,D_h)}{h^{(0)}} * \underset{(D_h,D_h)}{H} + \underset{(1,d)}{e^{(1)}} * \underset{(d,D_h)}{I} + \underset{(1,D_h)}{b_1} \right)$$

$$h^{(2)} = \sigma\left(\underset{(1,D_h)}{h^{(1)}} * \underset{(D_h,D_h)}{H} + \underset{(1,d)}{e^{(2)}} * \underset{(d,D_h)}{I} + \underset{(1,D_h)}{b_1} \right)$$

$$h^{(3)} = \sigma\left(\underset{(1,D_h)}{h^{(2)}} * \underset{(D_h,D_h)}{H} + \underset{(1,d)}{e^{(3)}} * \underset{(d,D_h)}{I} + \underset{(1,D_h)}{b_1} \right)$$

$$e^{(1)} = \left(\underset{(1,|V|)}{x^{(1)}} * \underset{(|V|,d)}{L} \right)$$

$$e^{(2)} = \left(\underset{(1,|V|)}{x^{(2)}} * \underset{(|V|,d)}{L} \right)$$

$$e^{(3)} = \left(\underset{(1,|V|)}{x^{(3)}} * \underset{(|V|,d)}{L} \right)$$