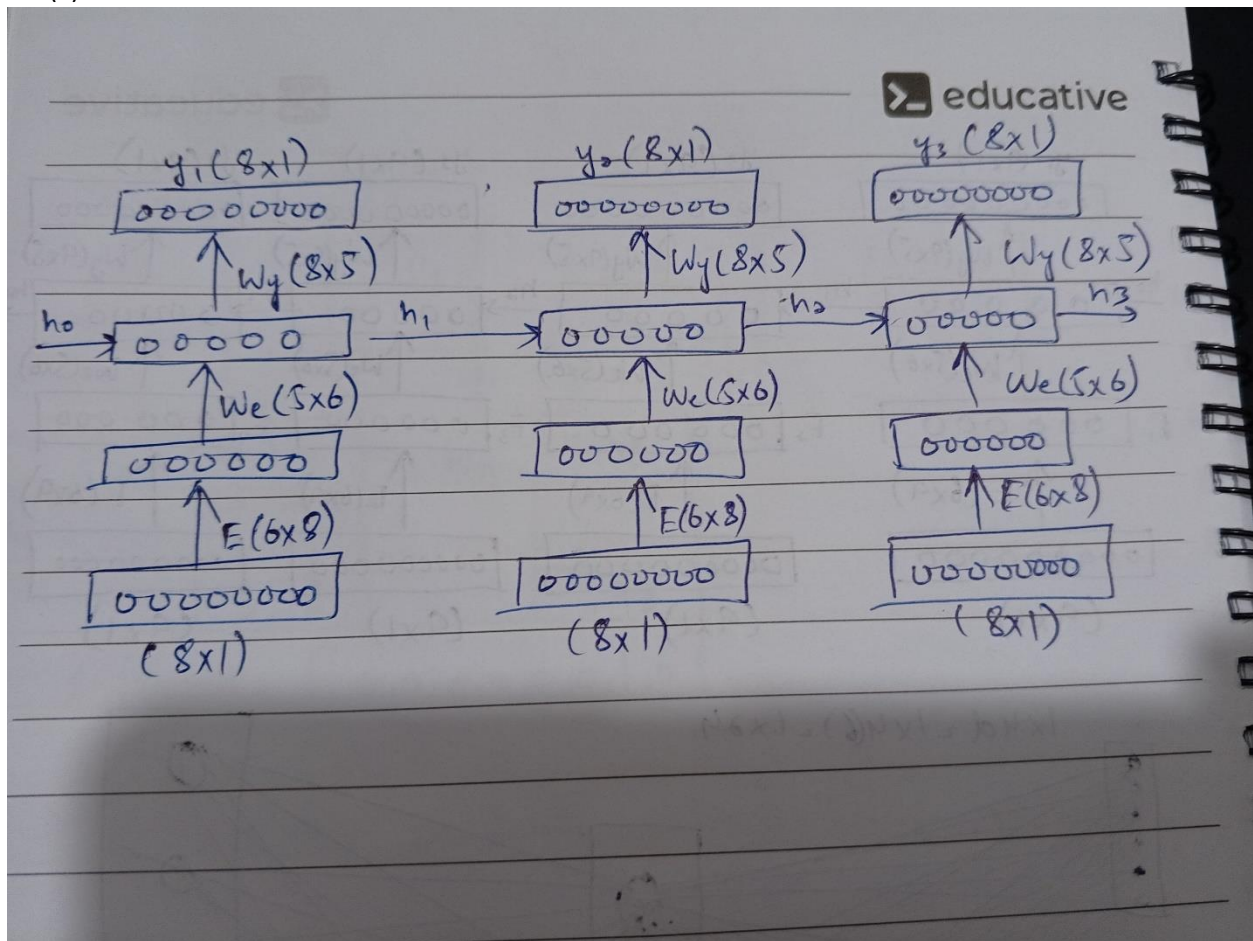


Q1 (a)



$$a_t = W_H h_{t-1} + W_X X_t \quad \text{Hidden Nodes}$$

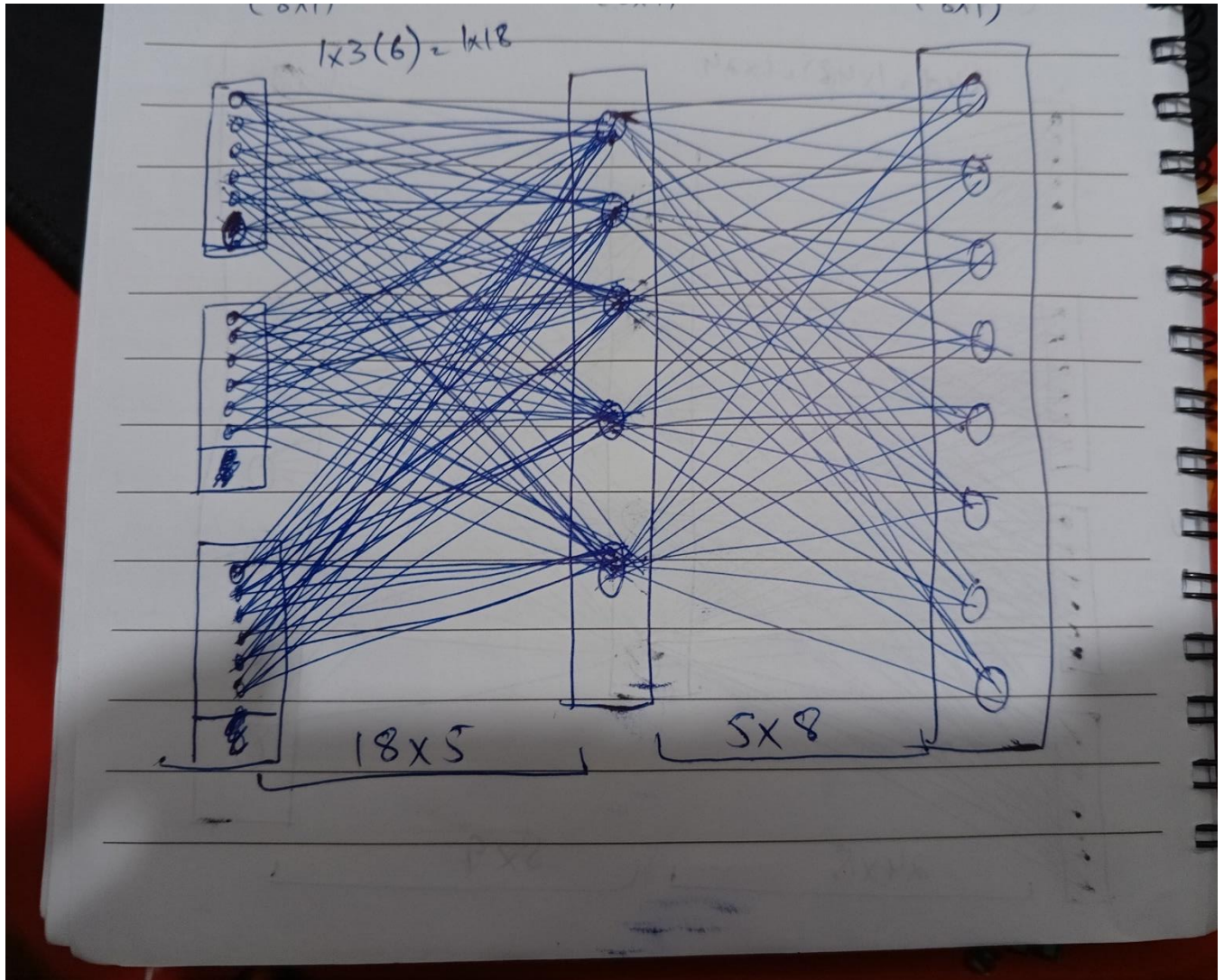
Output From Hidden State $\rightarrow h_t = \underbrace{\tanh(a_t)}_{\text{Hidden State}}$

Activation Function

Prediction at time $t \rightarrow y_t = \text{softmax}(W_Y h_t)$

(b) Total Number of Parameters= 4 (E, W_x, W_h, W_y)

Q2 (a)



$$e = (Ex_1, Ex_2, \dots, Ex)$$

$$h = \sigma(We + b)$$

$$z = Uh$$

$$y = \text{softmax}(z)$$

(b) Total Number of Parameters= 3 (E,W_x,W_y)