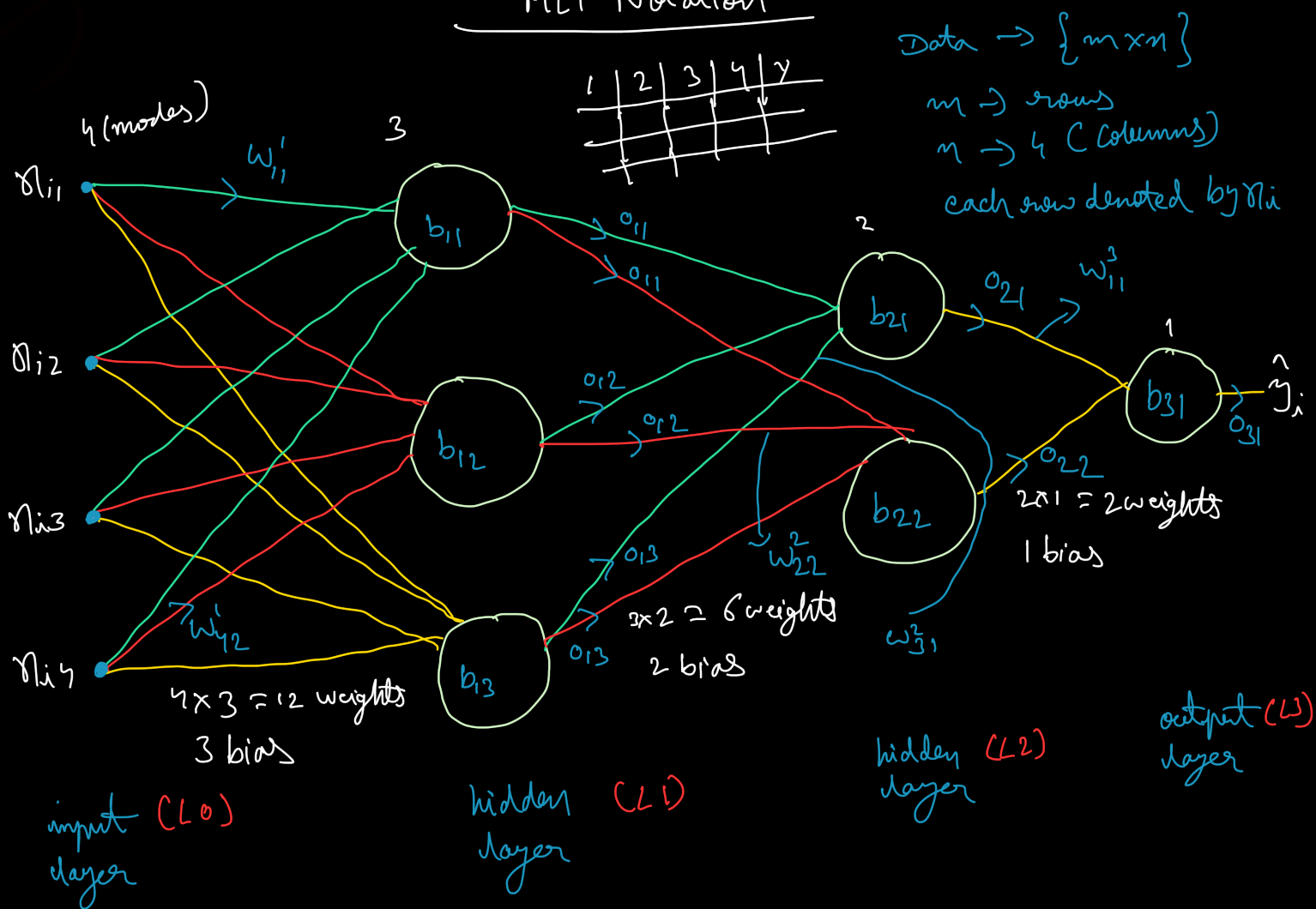


MLP Notation



Total trainable Parameters = $15 + 8 + 3 = 26$ (weights and Bias)

(w, b)

\rightarrow notation

Bias

b_{ij} : $i \rightarrow$ layer
 $j \rightarrow$ node

output

o_{ij} : $i \rightarrow$ layer
 $j \rightarrow$ node

weights

w^k_{ij} :

$k \rightarrow$ which layer it is going

$i \rightarrow$ coming from which node [Previous layer]

$j \rightarrow$ going to which node [in k]