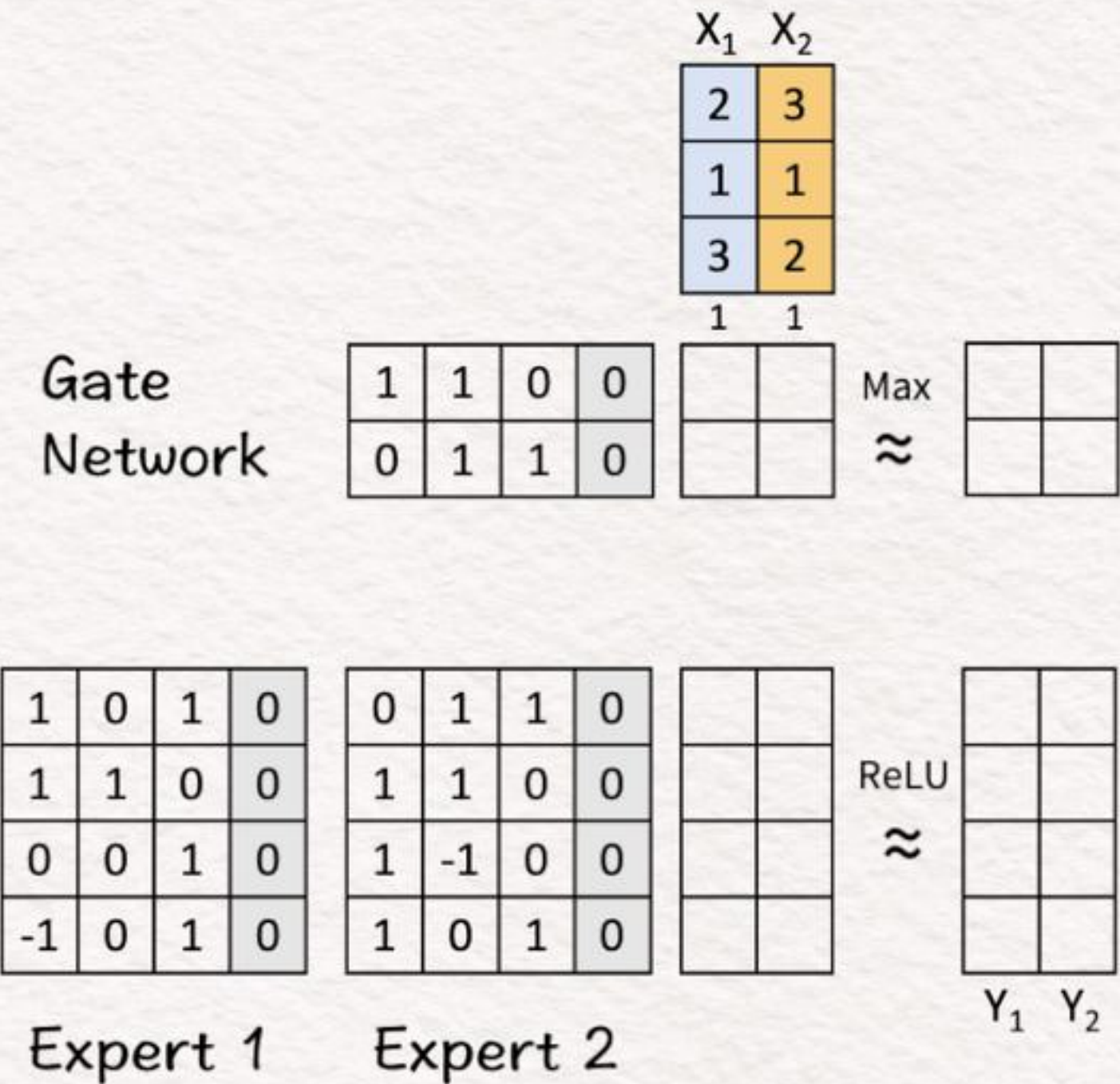


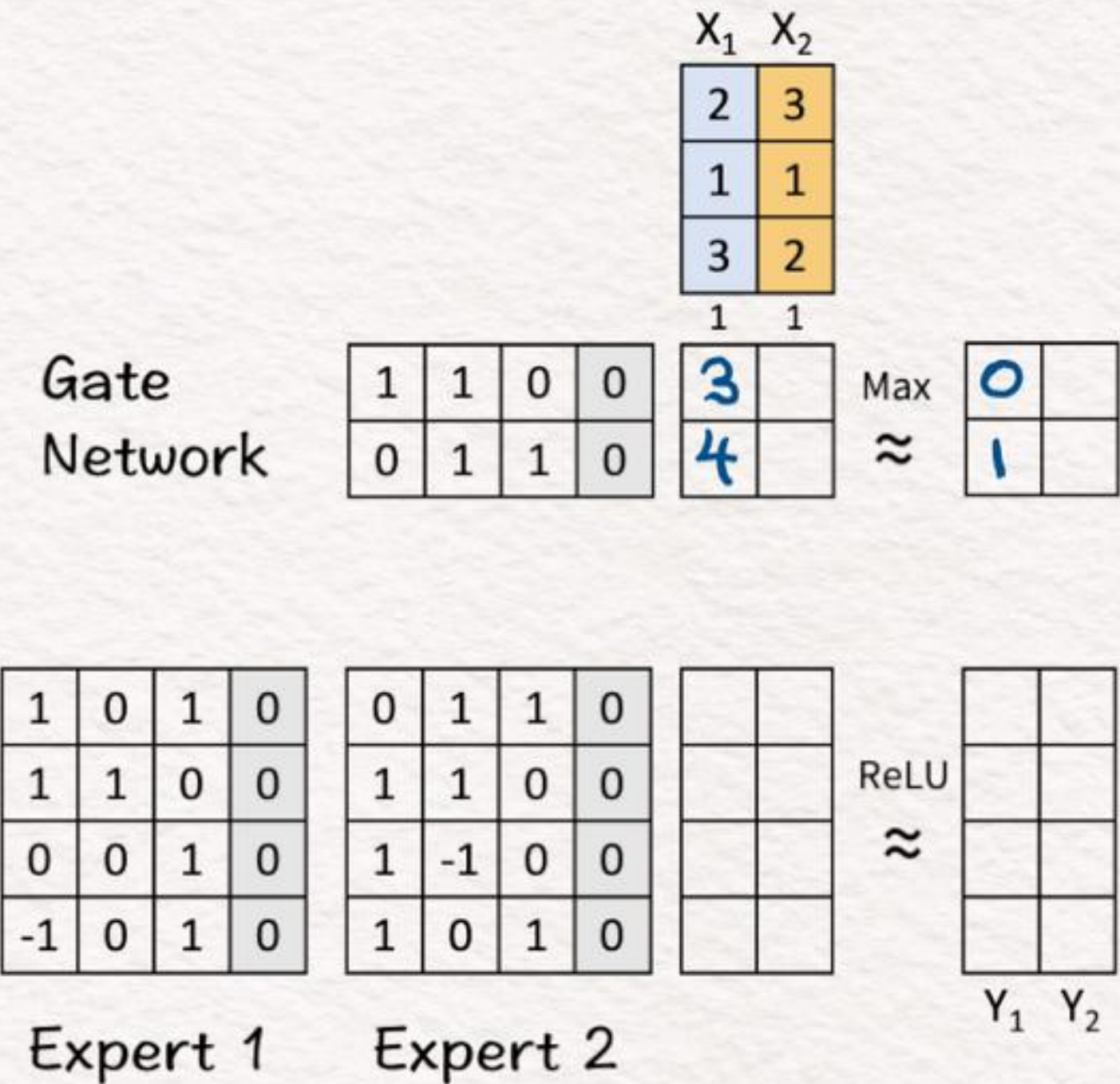
Mixture of Experts



© 2023 Tom Yeh

[1] [输入] MoE（混合专家）模块接收了两个令牌（蓝色、橙色）tokens (blue, orange). [1] [Inputs] The MoE block received two

Mixture of Experts

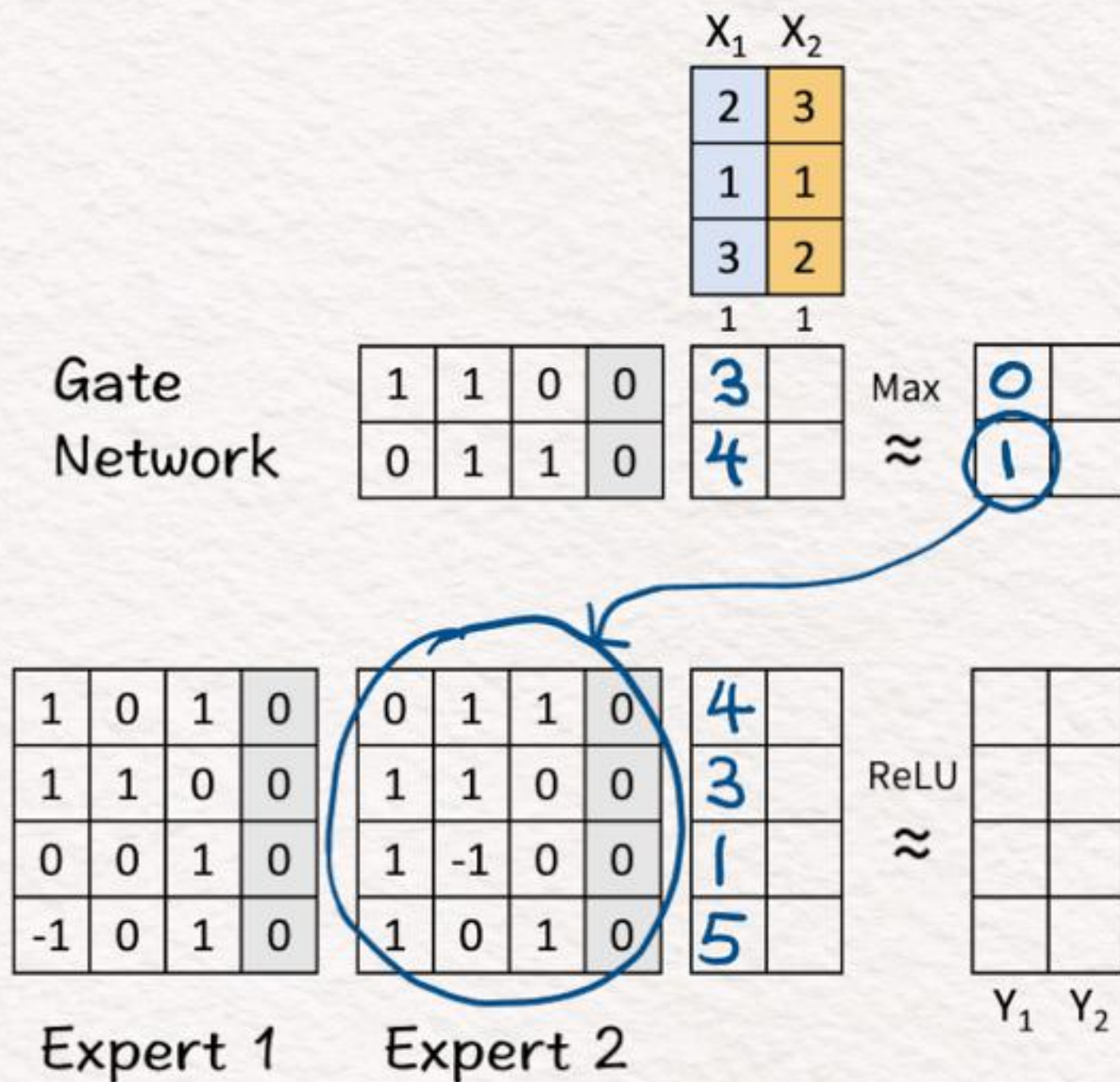


© 2023 Tom Yeh

[2] 门控网络处理了X1（蓝色）并确定应激活专家2。

[2] Gate Network processed X1 (blue) and determined Expert 2 should be activated.

Mixture of Experts

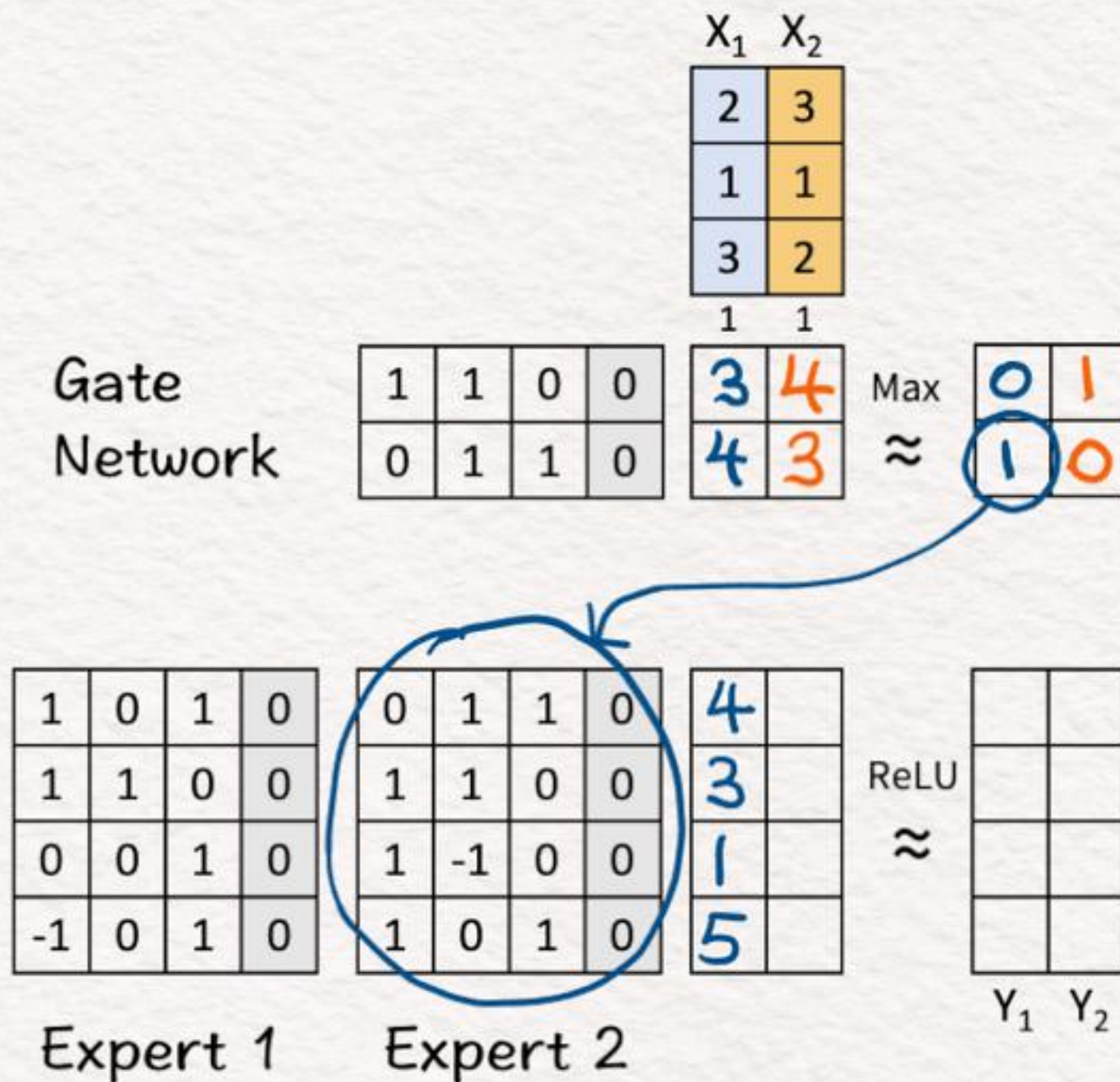


© 2023 Tom Yeh

[3] 专家2处理了 X_1 （蓝色）。

[3] Expert 2 processed X_1 (blue).

Mixture of Experts

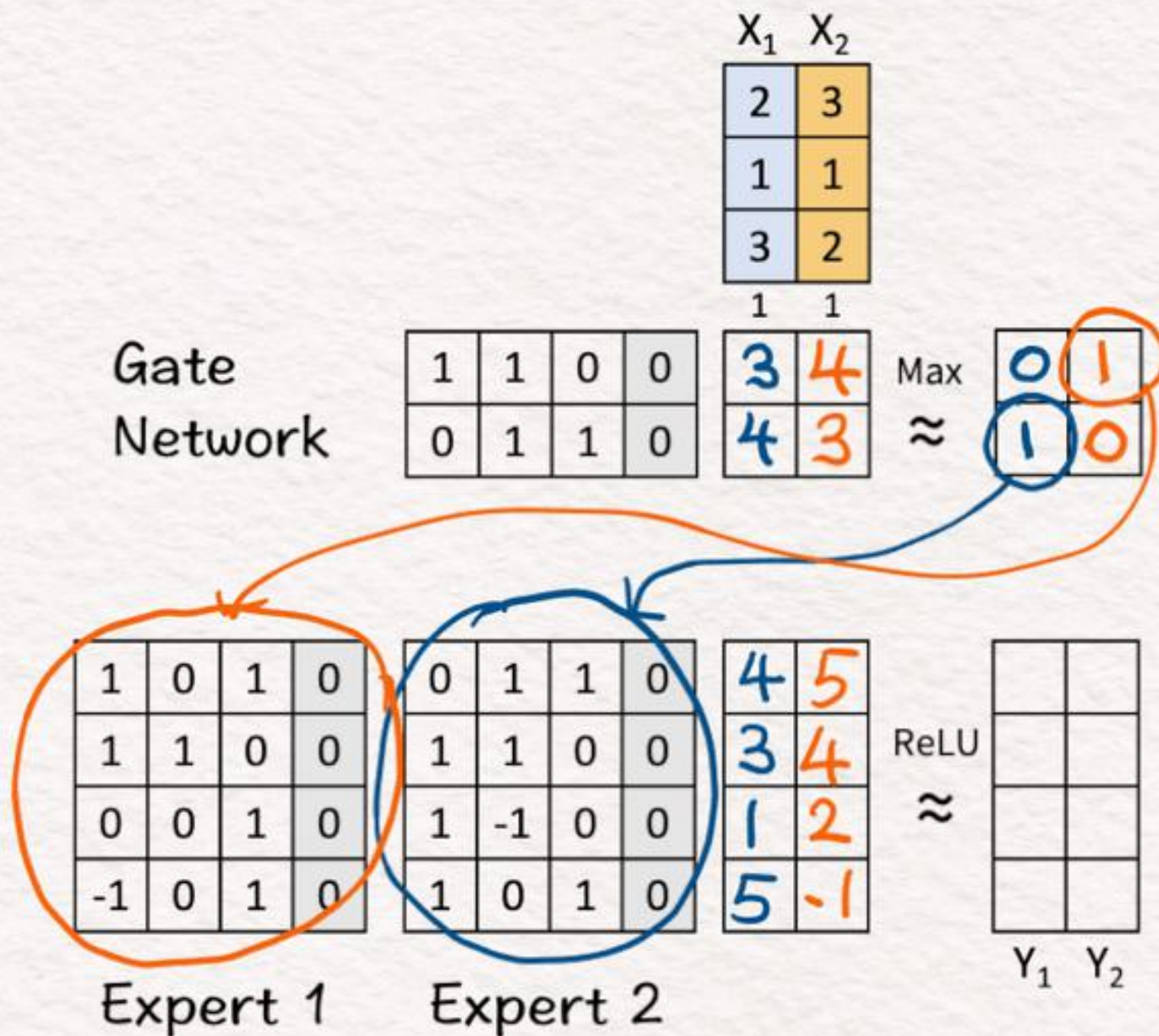


© 2023 Tom Yeh

[4]
门控网络处理了 X_2 （橙色）并确定应激活专家1。

[4] Gate Network processed X_2 (orange) and determined Expert 1 should be activated.

Mixture of Experts

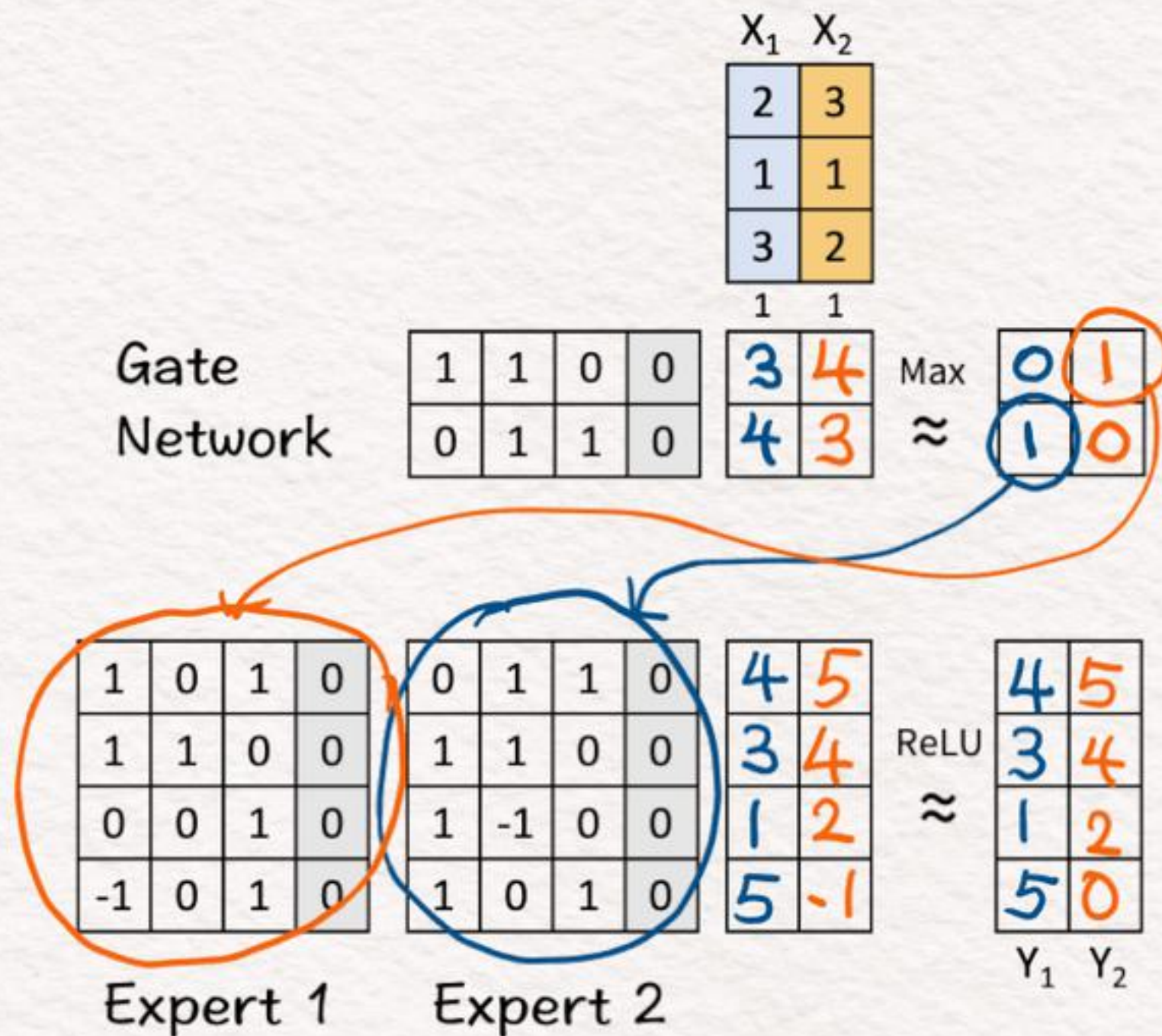


© 2023 Tom Yeh

[5] 专家1处理了 X_2 （橙色）。

[5] Expert 1 processed X_2 (orange).

Mixture of Experts



© 2023 Tom Yeh

[6] ReLU激活函数处理了专家的输出
并产生了最终输出。

[6] ReLU activation function processed the
outputs of the experts and produced the
final output.