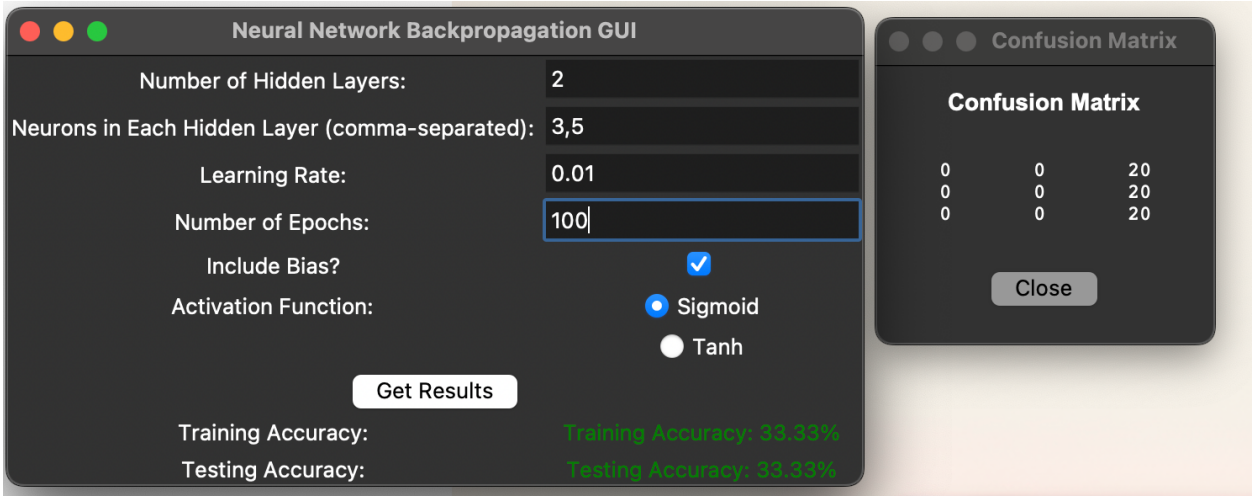
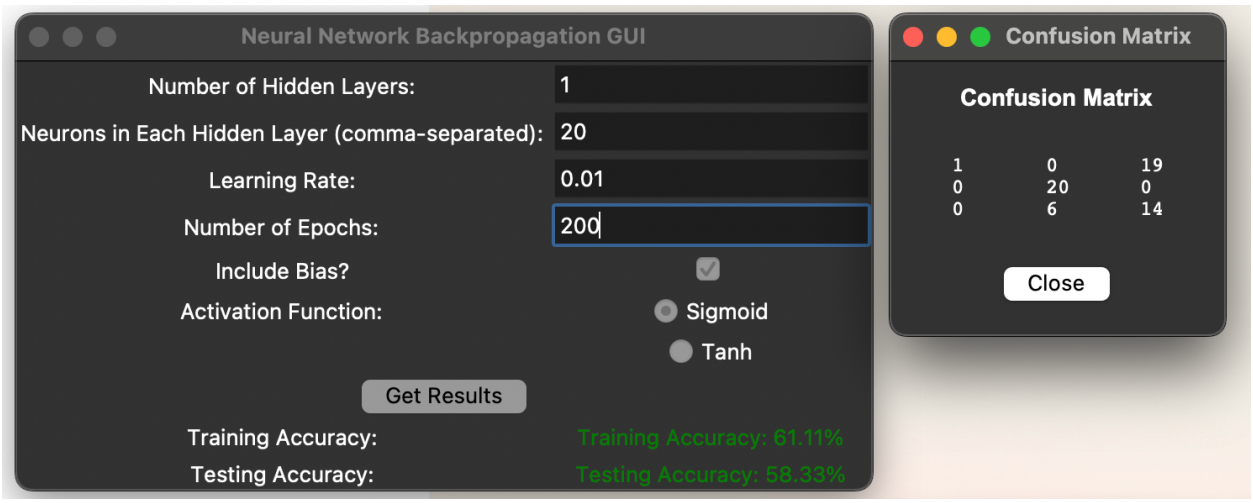
Neural Network Task 2
SC_3

No	Name	ID
1	حبيبة علاء الدين عادل الصناديدى	2021170161
2	مارك جمال انور فهمي	2021170423
3	ملك رأفت فؤاد فاضل	2021170538
4	رنا وحيد تمام عثمان	2021170195
5	مايكل نادر عماد الدين	2021170442
6	ماجد سيد محمود	2021170421

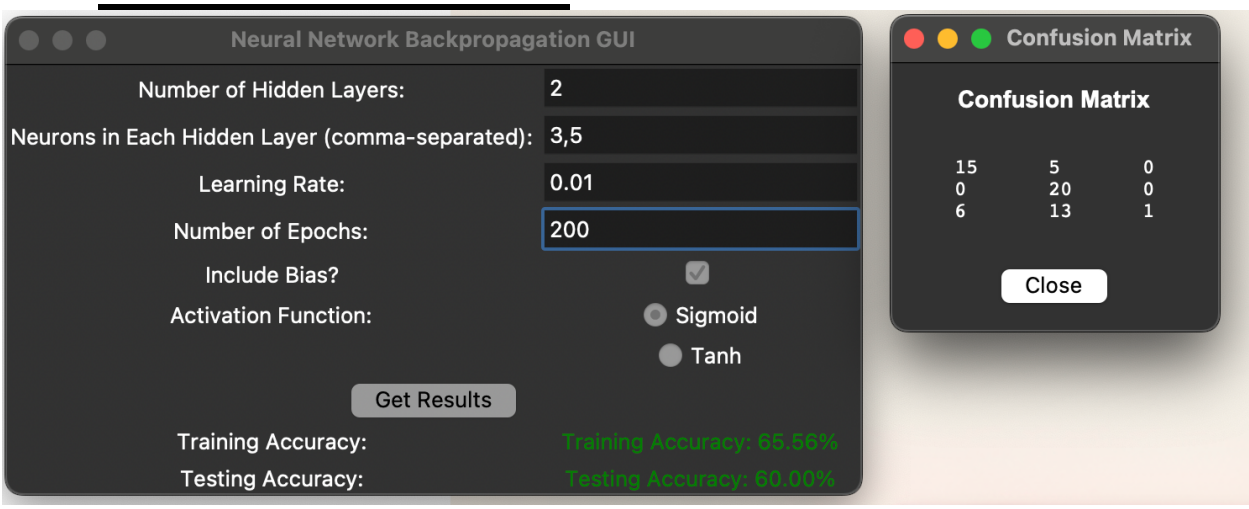
Sigmoid Function

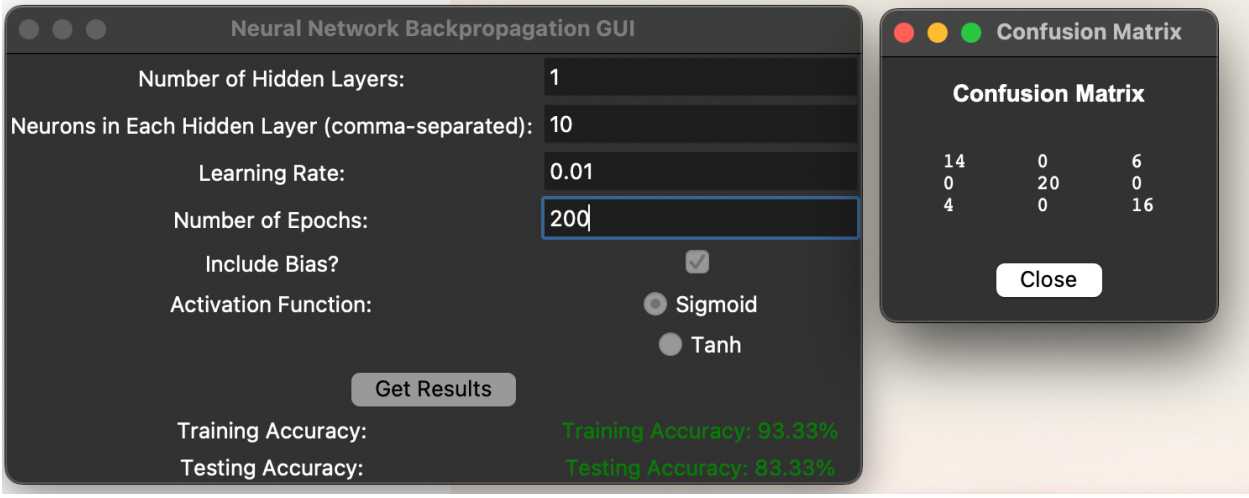
- **Init Random weights:**

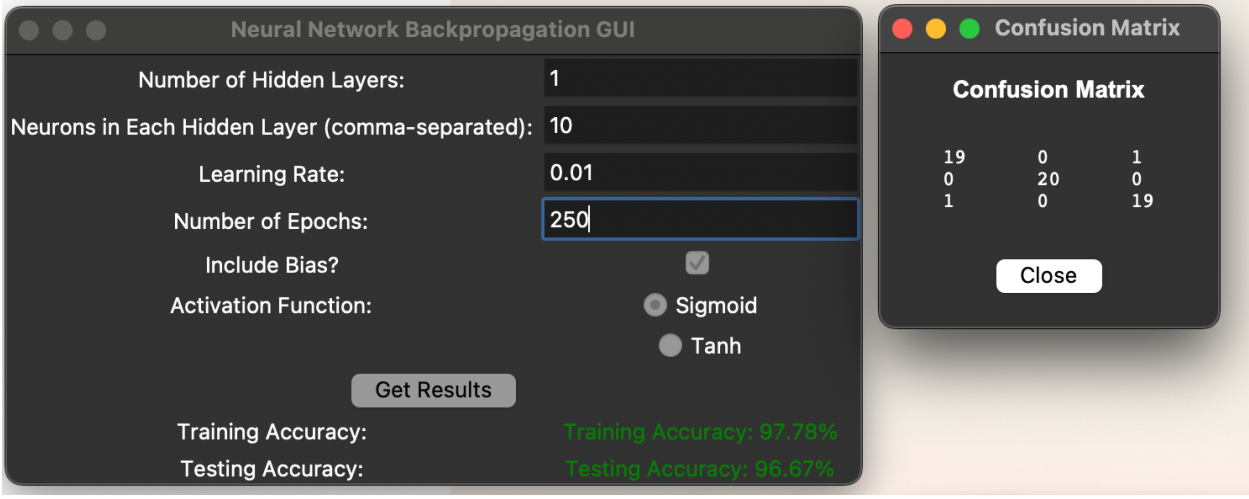




- **Xavier initialization:**







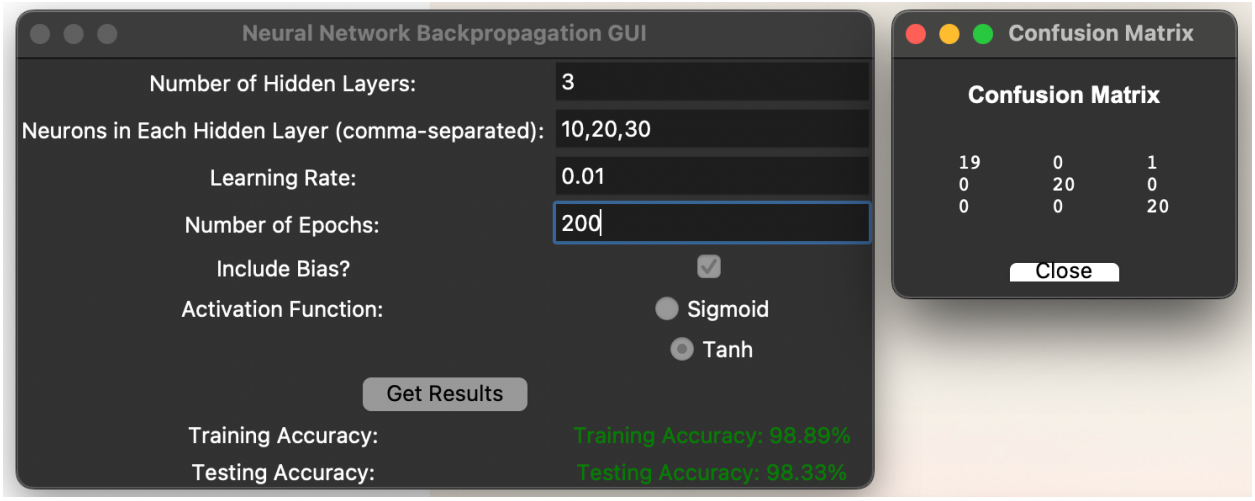
Conclusion:

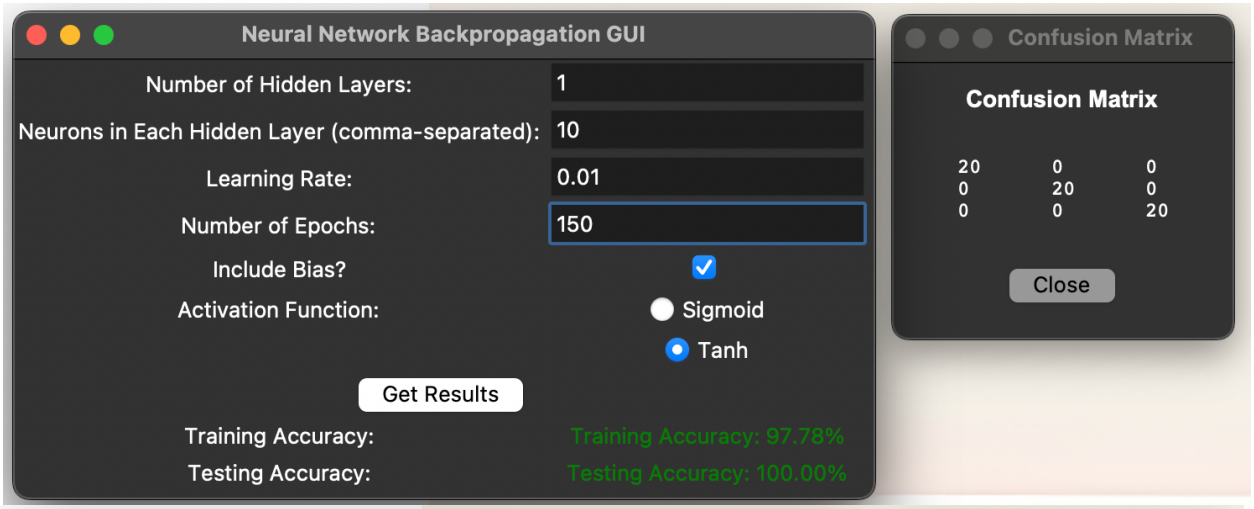
- Using random initialization of weights with the sigmoid activation function can lead to the vanishing gradient problem, particularly with multiple layers. While good results can still be achieved by increasing the number of epochs, this approach is not ideal due to the extended training time.

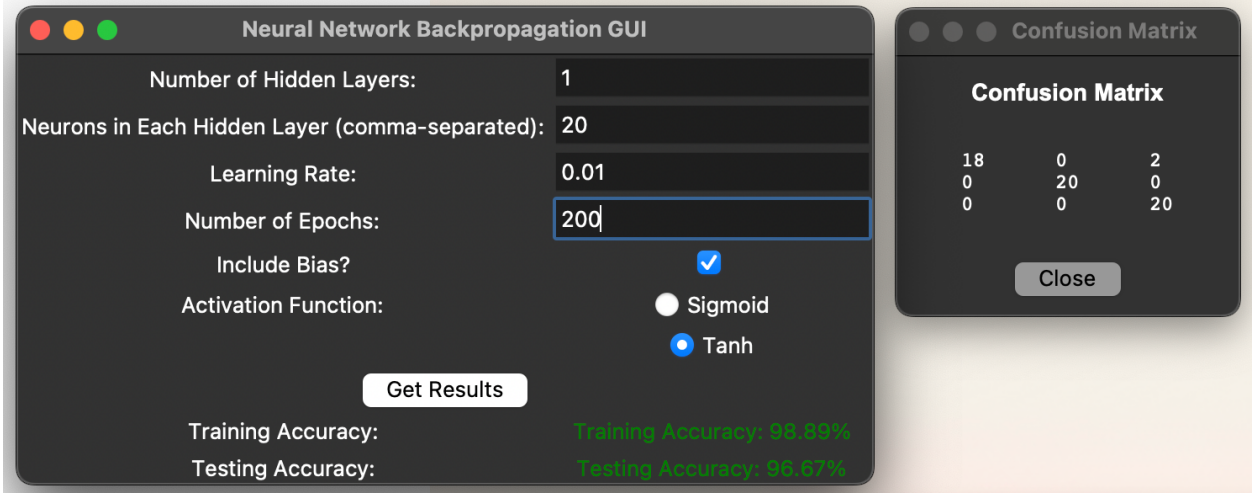
- On the other hand, Xavier initialization performs well regardless of the number of layers, neurons, or epochs. It effectively mitigates the vanishing gradient problem, enabling better results without the need for excessively long training periods.

Tanh

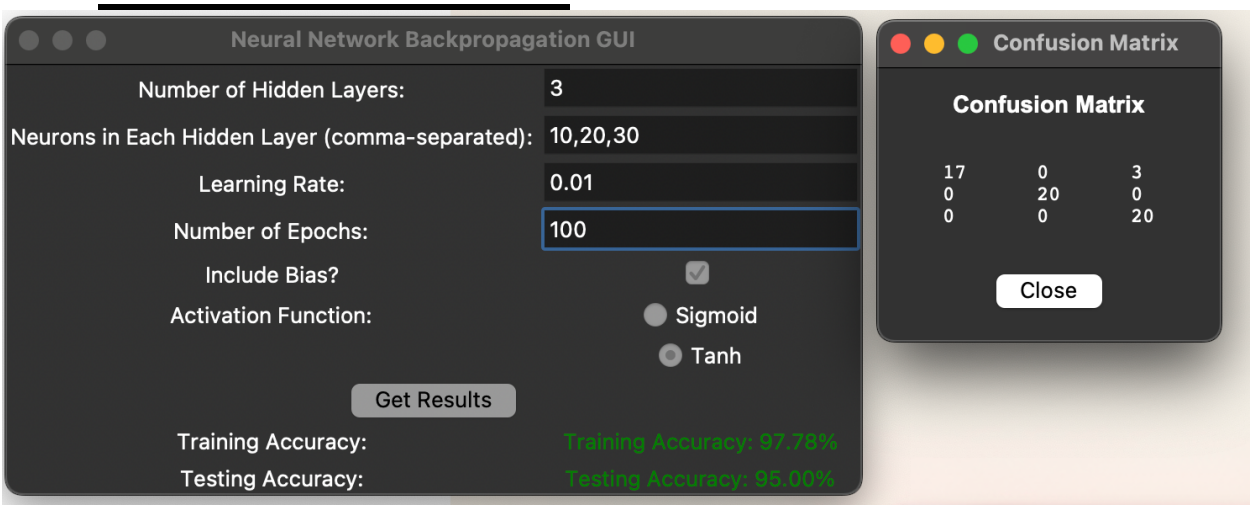
- **Init Random weights:**

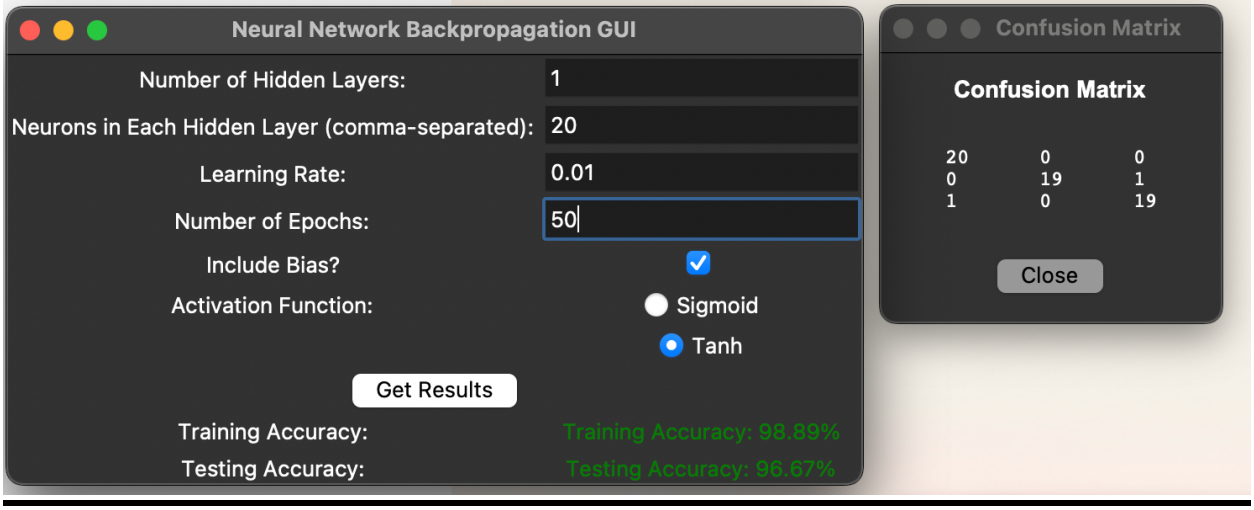






- **Xavier initialization:**





Conclusion:

For the tanh activation function, the accuracy does not significantly differ with different weight initialization methods. However, using Xavier initialization can still lead to higher accuracy with fewer epochs compared to random initialization, making the training process more efficient.