

Assignment Report

Observations

Upon completing the required instructions for the assignment, the following observations were noted:

Two Layer Network

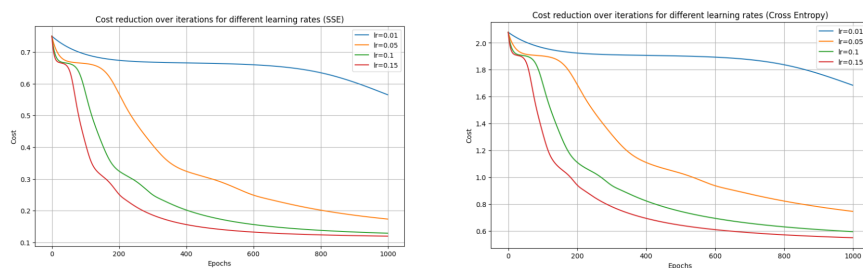


Figure 1: Two Layer Implementation showing SSE and CE loss functions.

- Based on the results in Figure 1, using the Sum of Squared Errors (SSE) as the loss function converged much faster than its Cross Entropy (CE) counterpart for corresponding learning rates.
- A learning rate of 0.15 achieved the least cost function during iteration.
- A learning rate of 0.2 showed some oscillations after initially converging.

Comparison of Loss Functions

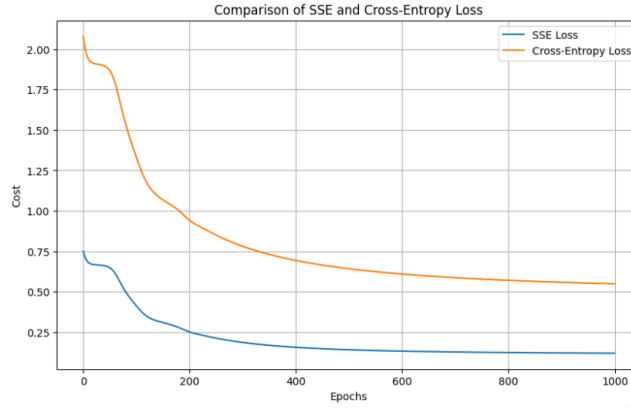


Figure 2: Comparison of Sum of Squared Errors (SSE) and Cross Entropy (CE) loss functions.

- The comparison in Figure 2 highlights the differences in convergence and accuracy between SSE and CE.

L Layer Network

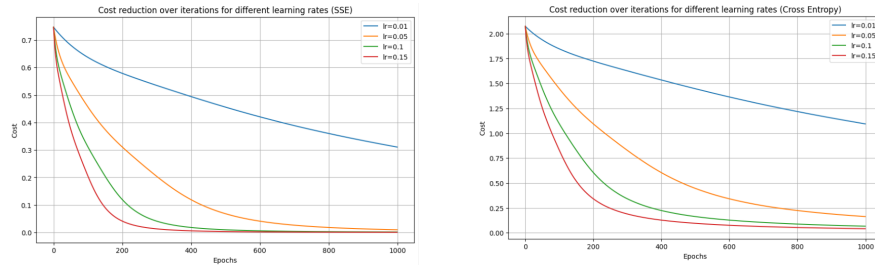


Figure 3: L Layer Implementation showing the effects of different learning rates.

- Different learning rates were tested, as shown in Figure 3. The behavior was similar to the two-layer model, especially without any fluctuations at the beginning of the iteration.

Keras Implementation

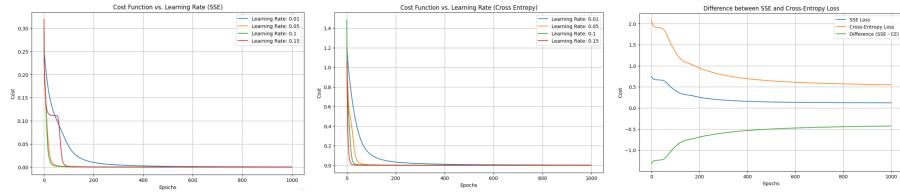


Figure 4: Keras Implementation for evaluating different learning rates and loss functions.

- As seen in Figure 4a, there is a change in the optimal learning rate. A value of 0.1 looks more effective when using the CE loss function, but remains 0.15 for the SSE loss function.
- The difference between the SSE and CE loss functions follows the same pattern as observed in the Two Layer Network, as seen in Figure 4c.