

A5

(a)

F1 Results:

- Test Accuracy: 0.9741
- Test Loss: 0.0885
- Total Parameters: 50,890

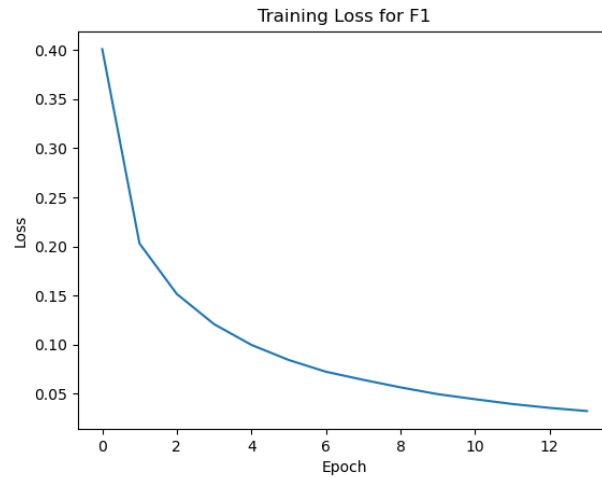


Figure 1: Training loss versus epoch for the F1 model (shallow and wide).

(b)

F2 Results:

- Test Accuracy: 0.9780
- Test Loss: 0.0767
- Total Parameters: 109,386

(c)

The F1 model, being shallow and wide, has fewer parameters (50,890) compared to the F2 model, which is deeper and narrower (109,386). Despite having fewer

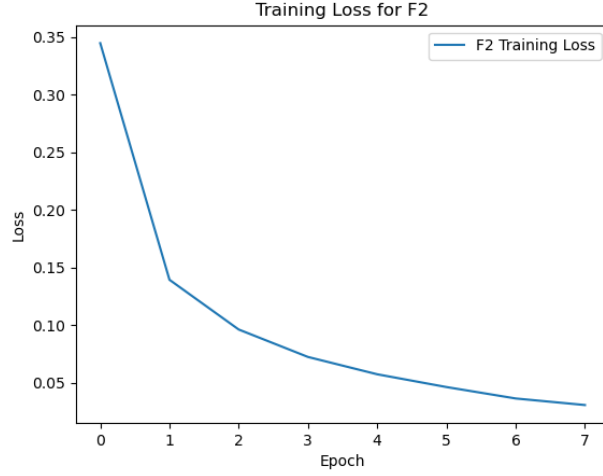


Figure 2: Training loss versus epoch for the F2 model (narrow and deep).

parameters, the F1 model achieves a high test accuracy of 0.9741, which is only slightly lower than the F2 model's test accuracy of 0.9780.

The F2 model's deeper architecture allows it to capture more complex patterns in the data, which likely contributes to its slightly better performance. However, this comes at the cost of a significantly higher number of parameters, which increases the computational and memory requirements.