A4

(b)

Cross-Entropy Loss: Figure **??** shows the training and validation loss curves for all models trained with the cross-entropy loss function.

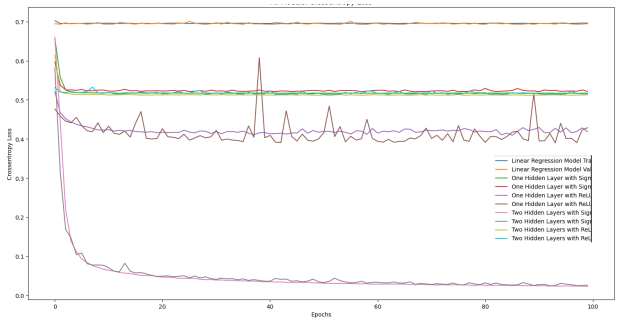

Figure 1: Training and Validation Loss Curves for All Models (Cross-Entropy Loss).

MSE Loss: Figure **??** shows the training and validation loss curves for all models trained with the MSE loss function.

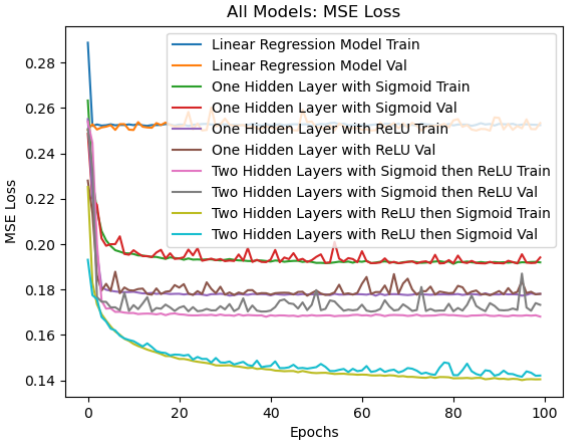

Figure 2: Training and Validation Loss Curves for All Models (MSE Loss).

(c)

Cross-Entropy Loss: The best model trained with cross-entropy loss was a network with two hidden layers using Sigmoid then ReLU activation functions.

It achieved a test set accuracy of 0.9928. The scatter plot of its predictions is shown in Figure **??**.

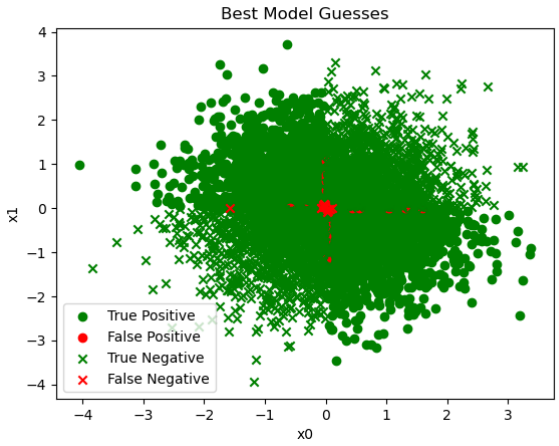

Figure 3: Scatter Plot of Best Cross-Entropy Model (Two Hidden Layers with Sigmoid then ReLU).

MSE Loss: The best model trained with MSE loss was a network with two hidden layers using ReLU then Sigmoid activation functions. It achieved a test set accuracy of 0.7212. The scatter plot of its predictions is shown in Figure **??**.

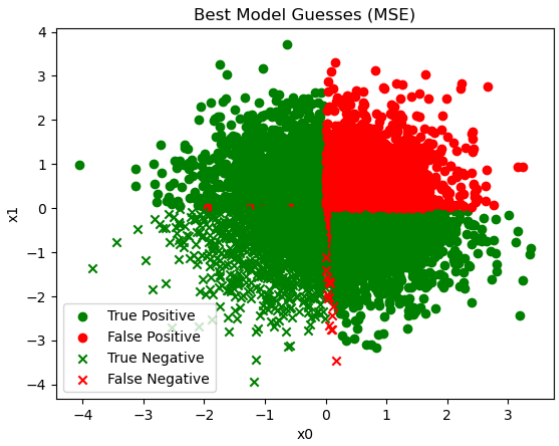

Figure 4: Scatter Plot of Best MSE Model (Two Hidden Layers with ReLU then Sigmoid).