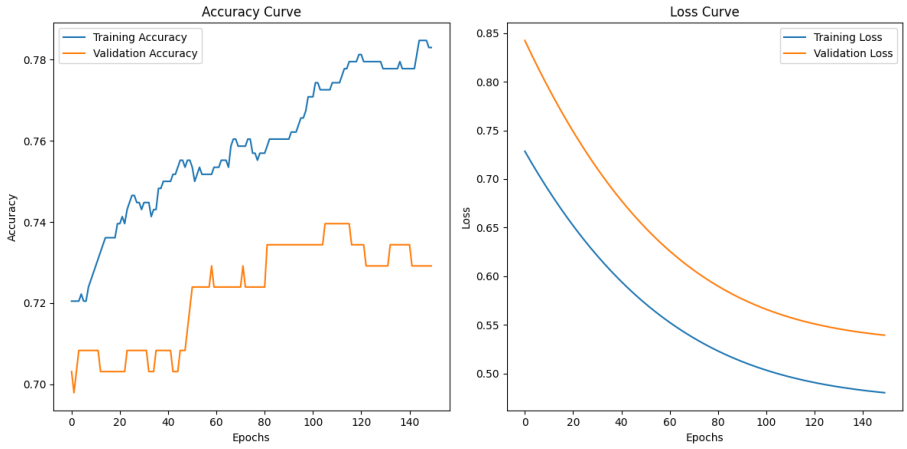

2.1

A simple FFNN achieved lower loss and higher training accuracy than LR because of nonlinear decision boundaries. Validation accuracy improved a bit as well. However the FFNN showed some signs of overfitting, unlike LR.

2.2

Increasing width improved training performance but eventually led to overfitting. Depth had moderate returns: 2 layers had small gain, but 3 layers slowed training and increased overfitting. Overall, wide networks helped more than deep ones on this dataset, but both made the model more sensitive to hyperparameters.

2.3

Setting the FFNN to 0 hidden layers and using the same hyperparameters made its loss and accuracy curves match LR exactly for all epochs.

3.

Logistic Regression: simple, stable, low overfitting, learns only linear boundaries. FFNN: more expressive and can outperform LR, but requires careful tuning and overfits more easily. When configured identically, FFNN reduces to LR and produces identical training curves.