HDS5210-W12
Wen-shan, Liu

1.

Data size	Configuration	Training error (1-trianing accuracy)	Validation error (1-val accuracy)	Time of execution(seconds)
1000	1 hidden layer 4 nodes	0.1375	0.145	0.1963
10000	1 hidden layer 4 nodes	0.0212	0.0155	1.6611
100000	1 hidden layer 4 nodes	0.0133	0.0129	10.0213
1000	2 hidden layers of 4 nodes each	0.2475	0.315	0.1597
10000	2 hidden layers of 4 nodes each	0.2438	0.242	0.9968
100000	2 hidden layers of 4 nodes each	0.0379	0.0373	5.7436

2. Based on the results, which model do you consider as superior, among the deep learning models fit?

In my experiment, I chose ReLU as the activation function, and the iteration is 500. Based on my result, the model with 1 hidden layer and 4 nodes is superior among the deep learning models fit. It achieves the lowest training error and validation error across all data sizes, indicating better learning and generalization performance.

Although the models with 2 hidden layers of 4 nodes each have slightly faster execution times, their training and validation errors are significantly higher, suggesting underfitting and poorer predictive accuracy.

3. Next, report the results (for the particular numbers of observations) from applying xgboost (week 11 – provide the relevant results here in a table). Comparing the results from XGBoost and deep learning models fit, which model would you say is superior to others? What is the basis for your judgment?

Model	Data size	Configuration	Training error (1-trianing accuracy)	Validation error (1-val accuracy)	Time of execution (seconds)
MLP Model1	1000	1 hidden layer 4 nodes	0.1375	0.145	0.1963
	10000	1 hidden layer 4 nodes	0.0212	0.0155	1.6611
	100000	1 hidden layer 4 nodes	0.0133	0.0129	10.0213
MLP Model2	1000	2 hidden layers of 4 nodes each	0.2475	0.315	0.1597
	10000	2 hidden layers of 4 nodes each	0.2438	0.242	0.9968
	100000	2 hidden layers of 4 nodes each	0.0379	0.0373	5.7436
XGBoost	1000	5 fold cross validation	0	0.04625	3.6489
	10000	5 fold cross validation	0	0.024125	3.208
	100000	5 fold cross validation	0.0007	0.011313	4.6264

Based on the results in the table, XGBoost is superior to the deep learning models (MLP) for all data sizes tested, XGBoost consistently has the lowest training and validation error across all sample sizes. Additionally, XGBoost's training error is almost zero, indicating excellent fit to the training data.

Even though XGBoost takes a little longer to execute than the most basic MLP models for small datasets, it is still very effective and quicker than the MLP with a single hidden layer for the biggest dataset. The main basis for my judgment is the validation error, as it best reflects the model’s generalization ability to unseen data. XGBoost not only achieves the lowest validation error but also maintains stable performance as the data size increases.