1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data size** | **Configuration** | **Training error** | **Validation error** | **Time of execution** |
| 1000 | 1 hidden layer 4 nodes | 0.005 | 0.0300 | 3.798 |
| 10000 | 1 hidden layer 4 nodes | 0.0025 | 0.0100 | 3.177 |
| 100000 | 1 hidden layer 4 nodes | 0.00100 | 0.00200 | 4.109 |
| 1000 | 2 hidden layers of 4 nodes each | 0.000625 | 0.00150 | 5.327 |
| 10000 | 2 hidden layers of 4 nodes each | 0.001112 | 0.00075 | 9.05 |
| 100000 | 2 hidden layers of 4 nodes each | 0.000850 | 0.00070 | 10.06 |

2.

The 2 hidden layers model with 4 nodes each trained on 10,000 data points demonstrates the best deep learning performance according to the results. The selected model configuration reaches the lowest validation error (0.00075) and demonstrates a training error (0.001112) that indicates strong generalization capabilities without overfitting. The 100,000 data point model with identical architecture achieves a lower validation error (0.00070) yet requires ten times more data and marginally longer execution time compared to the superior 10,000 data point model. The 10,000 data point model achieves the best combination of performance and data efficiency and computational cost which makes it the ideal solution for this problem.

3.

| **Method used** | **Dataset size** | **Testing-set predictive performance** | **Time taken for the model to be fit** |
| --- | --- | --- | --- |
| XGBoost in Python via scikit-learn and 5-fold CV | 1000 | 0.952546 | 0.376 |
| 10000 | 0.971261 | 0.804 |
| 100000 | 0.982619 | 3.452 |

All dataset sizes show better results when measured through XGBoost performance than deep learning models can achieve. The XGBoost model reaches 0.971261 testing-set predictive accuracy in 0.804 seconds when working with 10,000 observations while the best deep learning model (2 hidden layers) requires 9.05 seconds to achieve 0.00075 validation error. XGBoost delivers 0.952546 predictive performance at 0.376 seconds when operating on a dataset of 1,000 observations while performing better than deep learning models which experience higher validation errors.

The multiple evaluation metrics demonstrate XGBoost as the best model choice. These results show XGBoost demonstrates superior predictive capabilities at faster rates in all dataset conditions because it operates about ten times faster than deep learning models. The predictive capabilities of XGBoost improve substantially as data volume increases from 1,000 to 100,000 observations while its performance rating enhances from 0.952546 to 0.982619. By contrast the deep learning models demonstrate only modest performance improvements or sometimes decreased accuracy. This problem shows a better performance with gradient boosting because it uses effective feature interaction handling alongside built-in regularization mechanisms which prevent overfitting while working well with smaller data sizes.