Hammed Gafar

Sub-project 4

Method:

For this project, the Ionosphere dataset was used, and both the training and validation datasets were normalized using the normalization in utils.py in the code provided in the assignment. Different combinations of hyperparameters were used to train the logistic regression model, and the hyperparameter combinations that significantly reduced validation error were selected as the best.

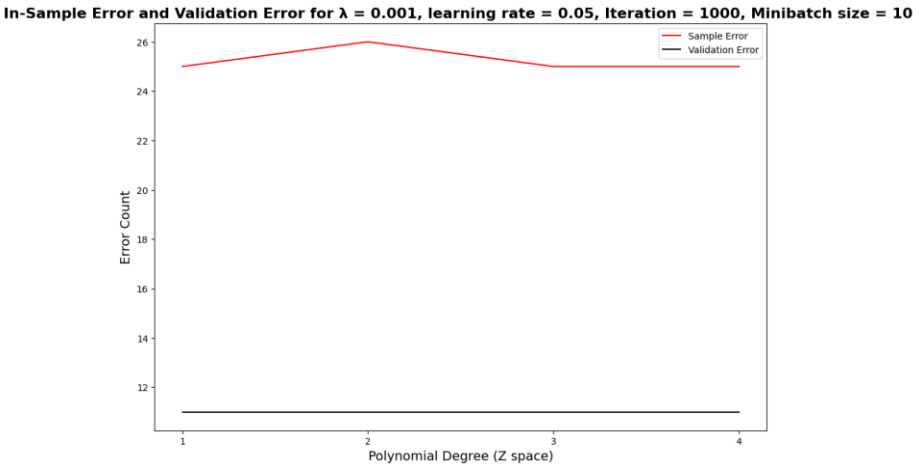
The hyperparameters are:

- The degree of Z-space.
- Lambda
- η (learning rate).
- The number of iterations.
- The size of the minibatch.

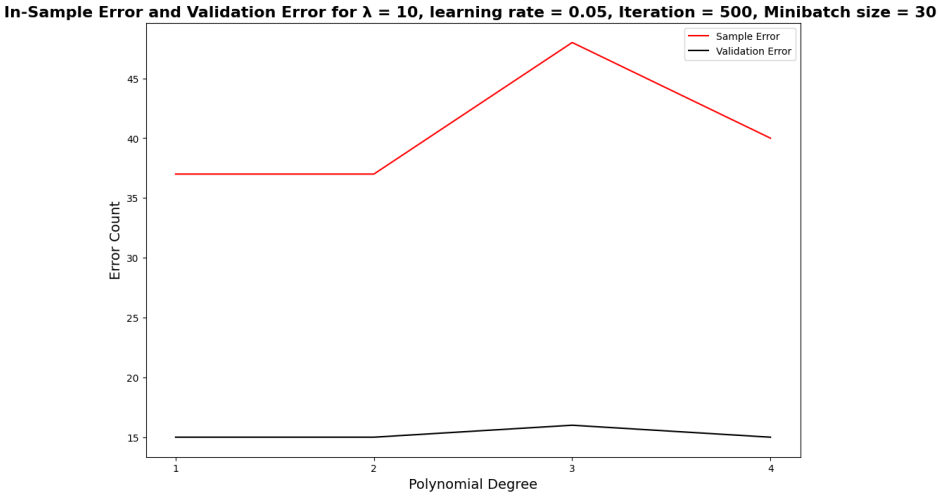
An automated script was used to test which hyperparameter combination significantly reduced the validation dataset error. The various hyperparameter combinations, sample errors, and validation errors were output in a CSV file for ease of visualization.

Various combinations values of the following parameters were used:

- Degree = 1, 2, 3, 4
- Lambda = 10, 1, 0.1, 0.01, 0.001
- Learning Rate = 0.1, 0.05, 0.01, 0.001
- Iteration = 500, 1000, 10000
- Mini batch size = 5, 10, 20, 30, 40

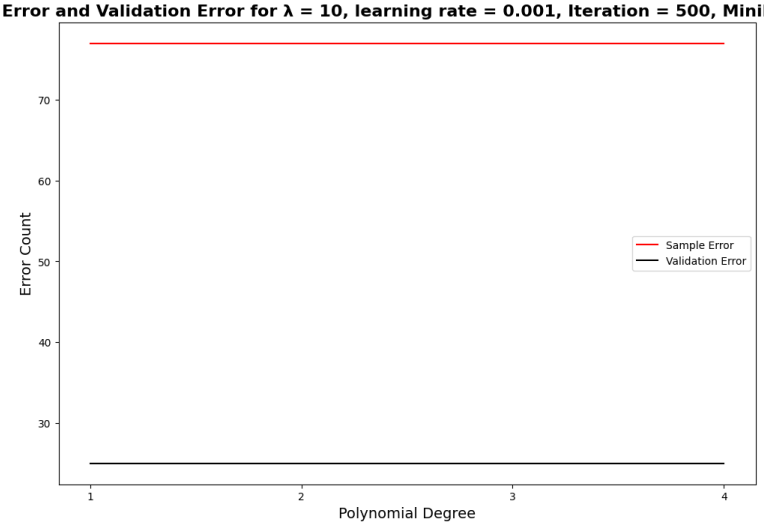


Using these parameters the validation error count was reduced to **10.** The learning rate of 0.05 and iteration of 1000 are quite large making it possible for the gradient descent to be close to the global minimum.



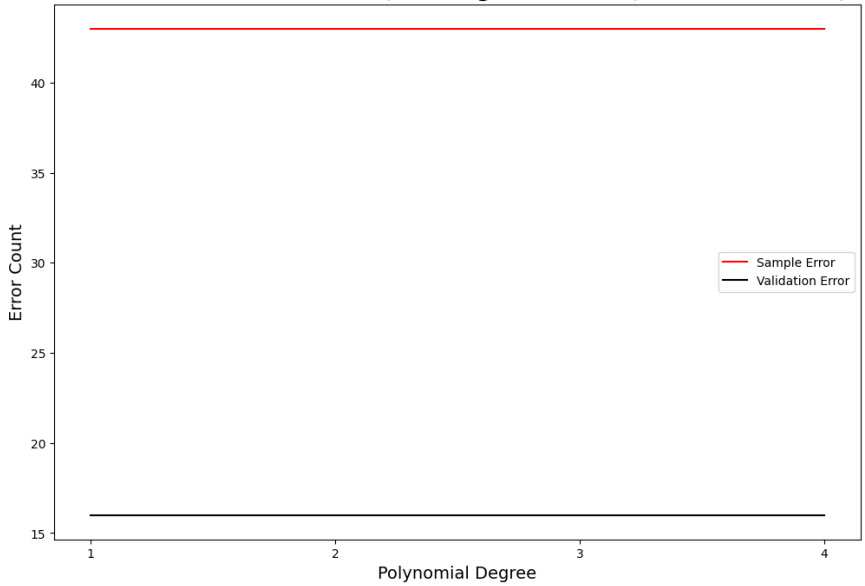
Using these parameters the validation error count was increased to **15.** The learning rate of 0.05 is quite fine and iteration of 500 is small compared to the previous diagram, thus, the gradient descent is less closer to the global minimum which explains why the error count increases by 5 compared to the previous figure.

In-Sample Error and Validation Error for λ = 10, learning rate = 0.001, Iteration = 500, Minibatch size = 10



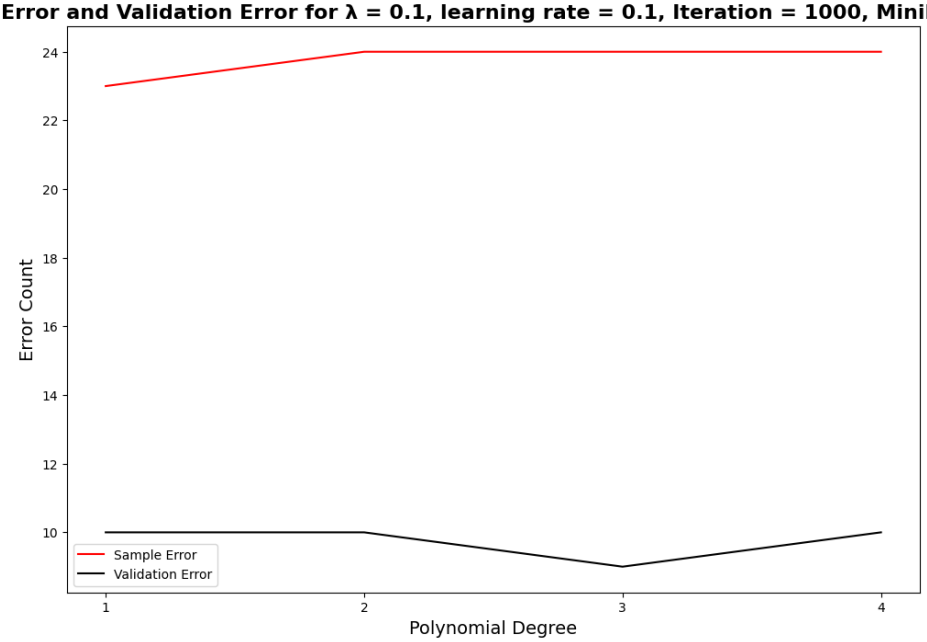
Using these parameters the validation error count is **25.** With a learning rate of 0.001 and 500 iterations, the algorithm is less likely to reach the global minimum which explains the high validation error count.

In-Sample Error and Validation Error for λ = 0.001, learning rate = 0.001, Iteration = 10000, Minibatch size = 40



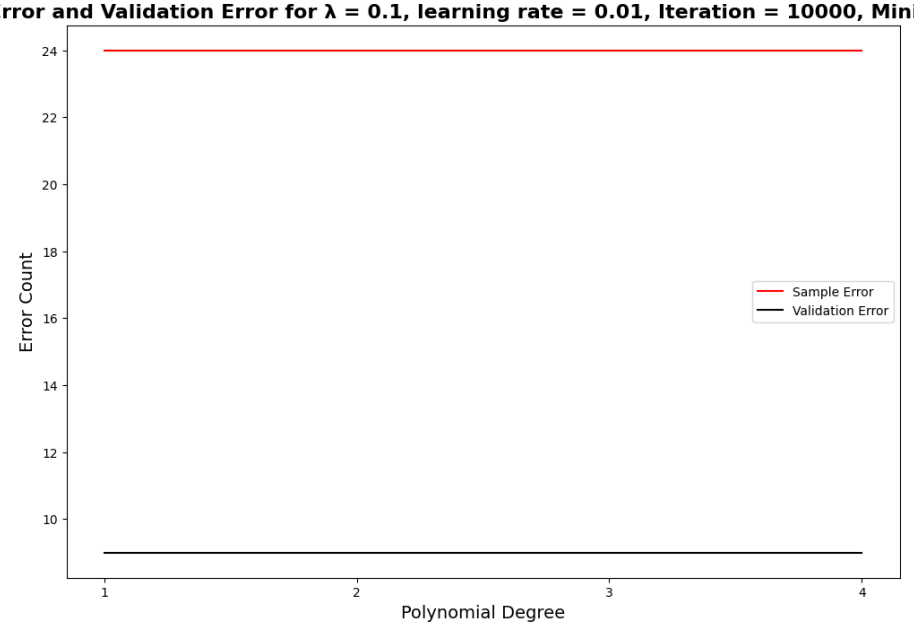
Using these parameters the validation error count is **16.** Due to the high number of iterations (10000) the algorithm performs better compared to the previous algorithm.

In-Sample Error and Validation Error for λ = 0.1, learning rate = 0.1, Iteration = 1000, Minibatch size = 20



Using these parameters the validation error count are **9 & 10** depending on the z-degree**.** This diagram suggests that taking a longer walk (0.1 learning rate) and a reasonable amount of iteration will further reduce the error.

In-Sample Error and Validation Error for λ = 0.1, learning rate = 0.01, Iteration = 10000, Minibatch size = 20



This diagram further emphasizes the suggestion of the previous diagram, longer walk (higher learning rate) and more iterations reduce the validation error significantly. Using

these combinations of hyperparameters brings down the validation error count to a record of **8.** Therefore, this is the best hyperparameter combination so far.

Conclusion

The best hyperparameter parameter combinations are:

λ = 0.1, learning rate = 0.1, Iteration = 10000, Minibatch size = 20, Z-degree= (1, 2, 3, 4).

The z-degree does not have a significant effect on the validation error due to low **λ of 0.1.** In addition, the minibatch size does not have a significant effect because the data was shuffled, thus, each minibatch is a representative of the entire dataset.