NYU Tandon Machine Learning

Mini Project

Designed by: Arinola Soyemi

The Model And Loss Function

The Model:

$$f(Xi;w) = w0 + w1(xi1) + w2(xi2) + w3(xi3) + w4(xi4) + w5(xi5) + w6(xi1)^2 + w7(xi2)^2 + w8(xi3)^2 + w9(xi4)^2 + w10(xi5)^2 + w11(xi1)^3 + w12(xi2)^3 + w13(xi3)^3 + w14(xi4)^3 + w15(xi5)^3$$

Cost Function: MSE + lambda(sum of weights squared)

Weights

w 0 = [0.]

w 1 = [0.00305917]

w 2 = [-0.00035214]

w 3 = [-0.00064732]

w 4 = [0.02637792]

w 5 = [0.39669961]

w 6 = [0.00305917]

w 7 = [-0.00035214]

w 8 = [-0.00064732]

w 9 = [0.02637792]

w 10 = [0.39669961]

w 11 = [0.00305917]

w 12 = [-0.00035214]

w 13 = [-0.00064732]

w 14 = [0.02637792]

w 15 = [0.39669961]

Ridge Regression and Best Alpha

I chose to use a ridge regression because each of the features were necessary to determine the weight of the fishes.

The optimal alpha we found for the ridge regression was ~0.69, which gave a validation RMSE of ~ 46.114.

Model Performance On Training and Validation Set

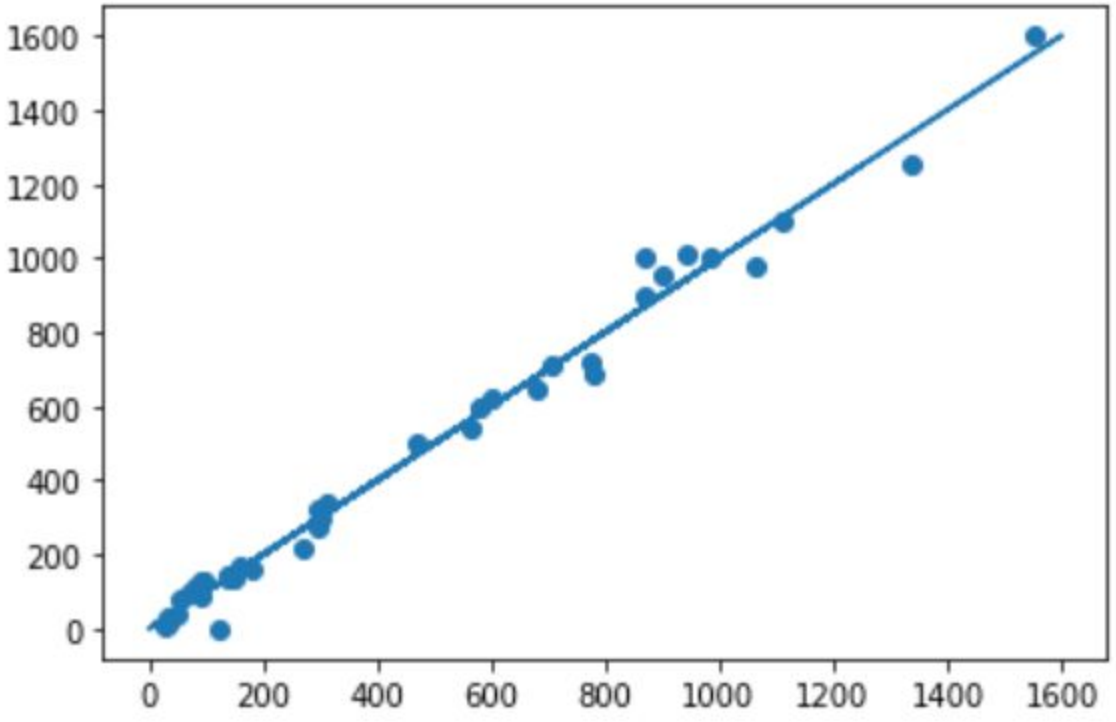
Mean Squared Error of Training Set:
1673.1883791104574

Mean Squared Error of Validation Set:
2126.5394072694244

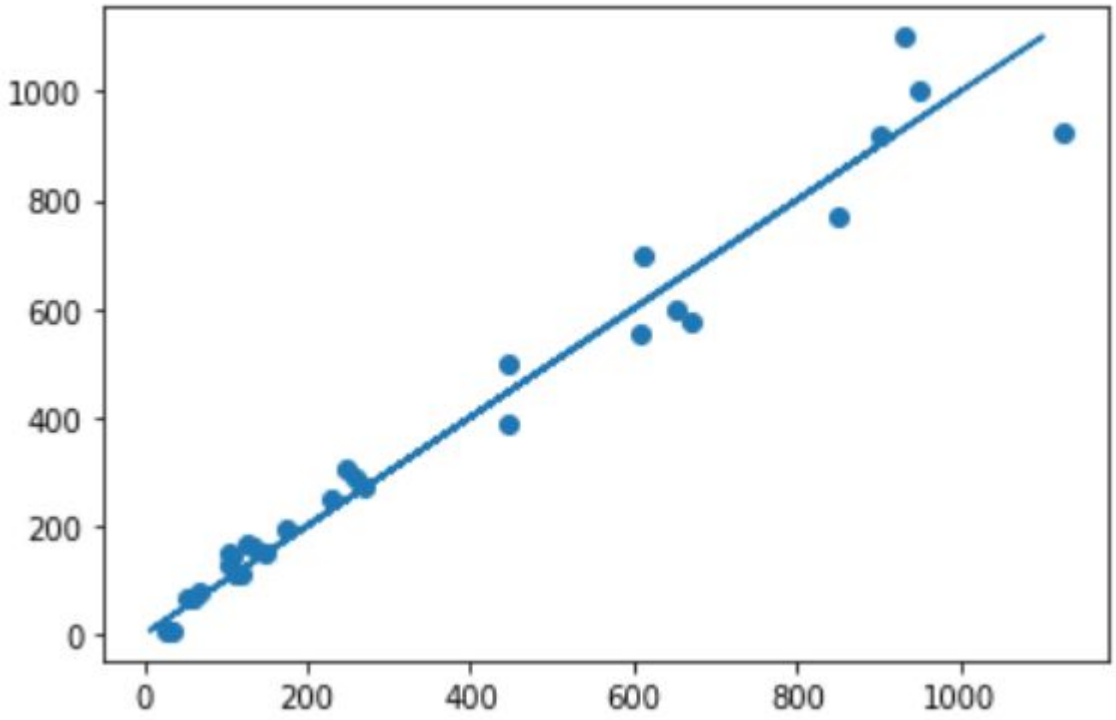
Mean Absolute Error of Training Set:
26.820291225131243

Mean Absolute Error of Validation Set:
33.85769272601199

VALIDATION ERROR VISUALIZATION



TEST ERROR VISUALIZATION



Model PERformance Of Test Set

Test MSE:4033.260819832647

Test MAE:33.85768114036656

Conclusion

1. First, I found the model. Then I found the cost function by using this formula:MSE + lambda(sum of weights squared)
2. Since I needed each feature to determine the weight of the fishes, I decided to use ridge regression. Our results: Ridge regression: ~0.69, optimal alpha:~ 46.114.
3. I found the Mean Squared Error/ Mean Absolute Error of the training set and validation set.
4. Results for the model for performance of test set: MSE: 4033.260819832647 MAE:33.85768114036656