REPORT

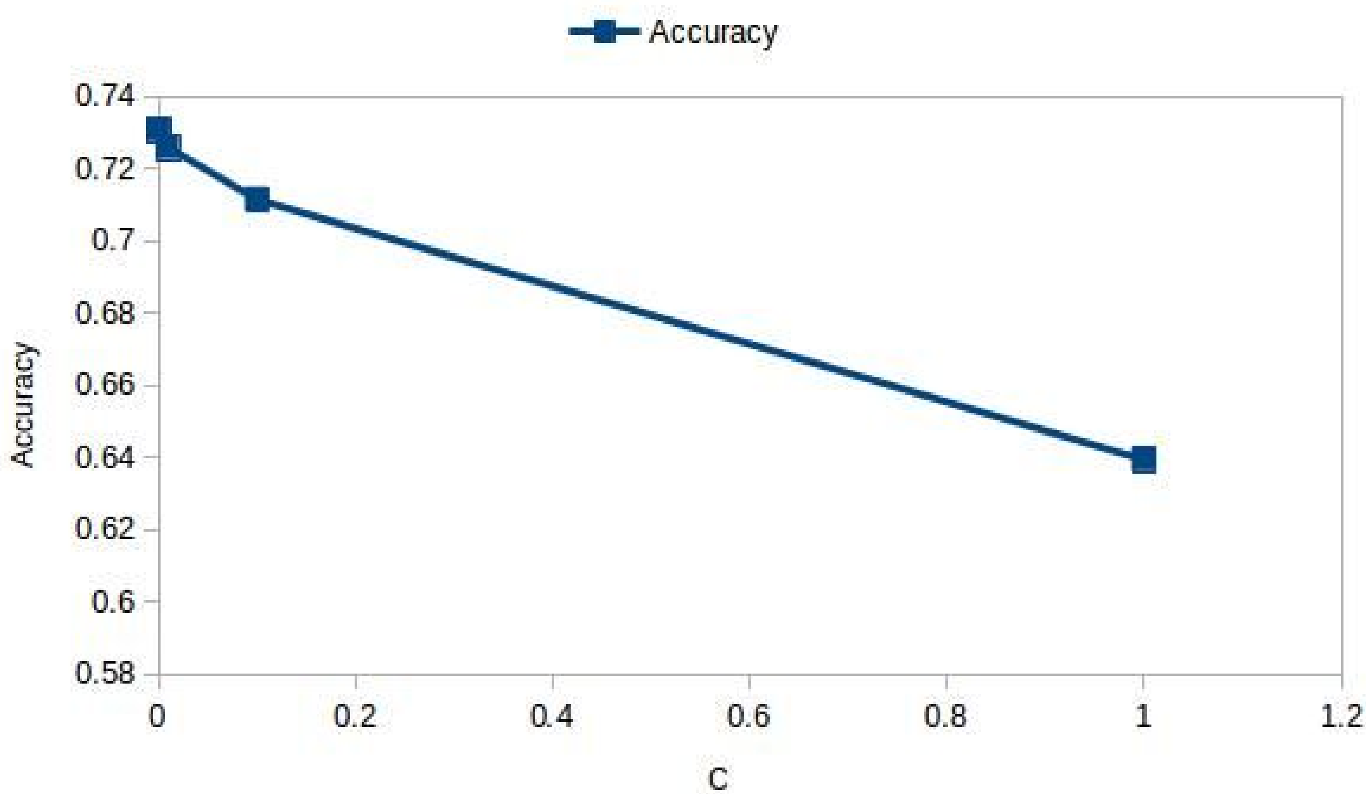
Shyam Nandan
20172088

a) Lasso (L1)

The lasso estimate thus solves the minimization of the least-squares penalty with $\alpha||w||_1$ added, where α is a constant and $||w||_1$ is the ℓ_1-norm of the parameter vector. Now in this α Is used as hyperparameter to validate the accuracy.

Best Accuracy : α = 1e-4 , Accuracy = 0.730769230769

Plot Showing Accuracy(Y-axis) vs Alpha(X-axis)

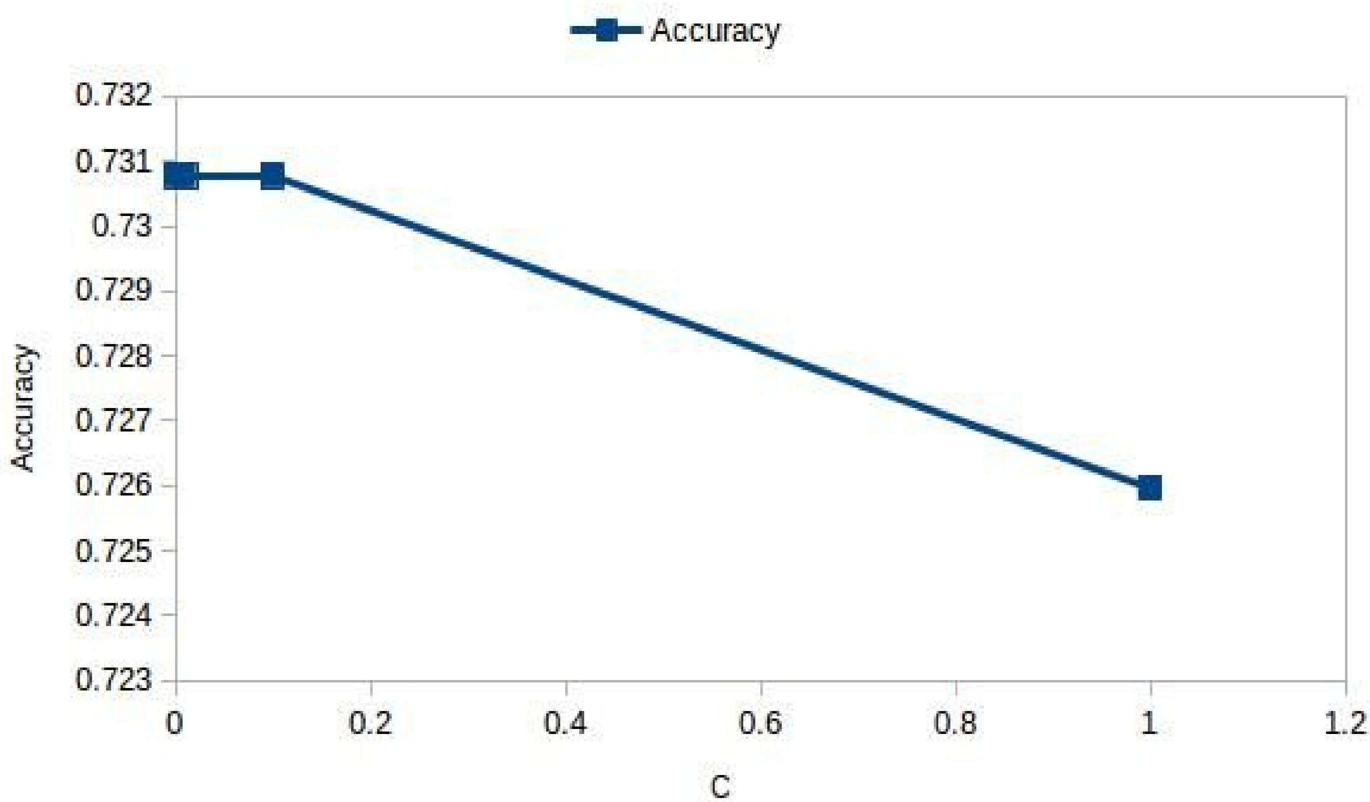


b)Ridge (L2)

Here in Ridge, $\alpha \geq 0$ is a complexity parameter that controls the amount of shrinkage: the larger the value of α, the greater the amount of shrinkage and thus the coefficients become more robust to collinearity. So this parameter was chosen to be hyperparameter.

Best Accuracy : α **= 1e-3 , Accuracy = 0.732**

<u>**Plot Showing Accuracy(Y-axis) vs Alpha(X-axis)**</u>

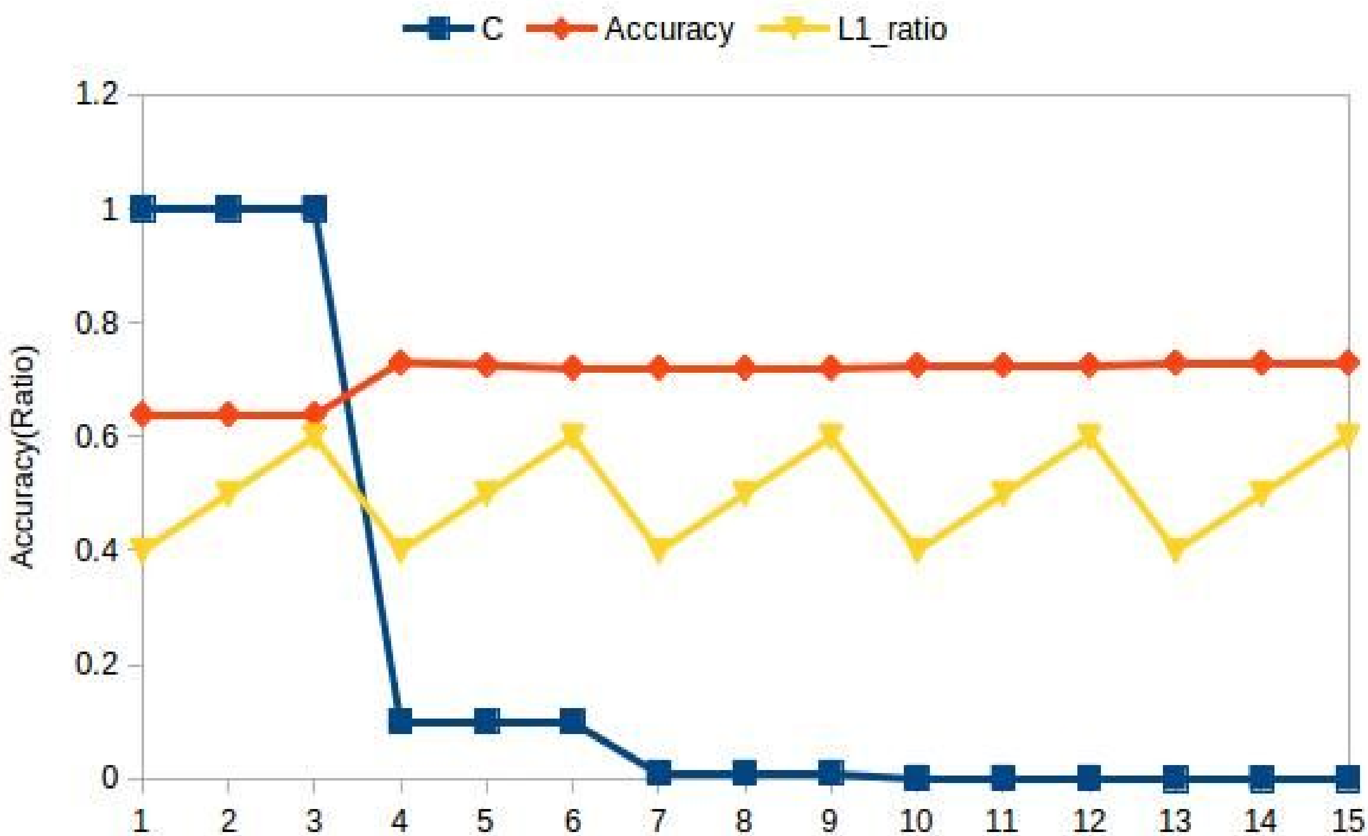


(c) Elastic net (Lasso and Ridge combined)

Elastic Net is a linear regression model trained with L1 and L2 prior as regularizer. This combination allows for learning a sparse model where few of the weights are non-zero like Lasso, while still maintaining the regularization properties of Ridge. We control the convex combination of L1 and L2 using the l1_ratio parameter also parameters alpha (α).

Best Accuracy : α = 1e-3 , Accuracy = 0.733, L1_ratio = 0.4

<ins>**Plot Showing Accuracy vs Alpha vs L1_ratio**</ins>



d) No Regularization

Coefficient estimates for Ordinary Least Squares rely on the independence of the model terms i.e they do not depend on any hyperparameter.

Best Accuracy : Accuracy = 0.731