REPORT

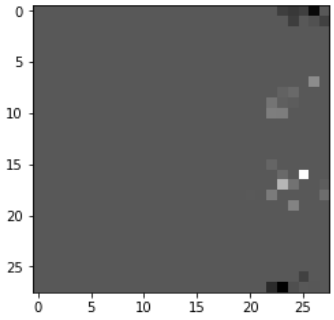
Shyam Nandan
20172088

λ is the regularization parameter for the logistic regression but sklearn Inverse of regularization strength ‘C’ . So, the below plots are produced with different value C.

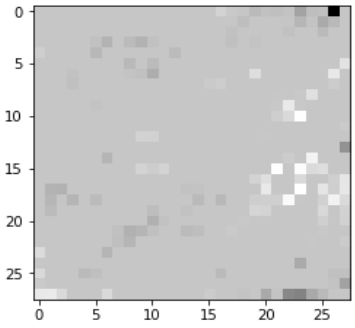
L1 loss

Accuracy on test data with L1 penalty: 95.20%
[Parameters - tolerance = 0.1, C = 10]

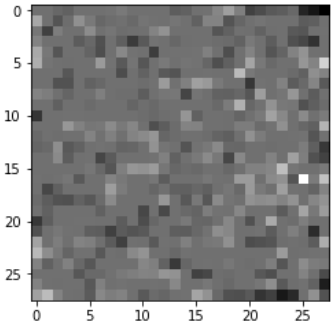
C = 0.01



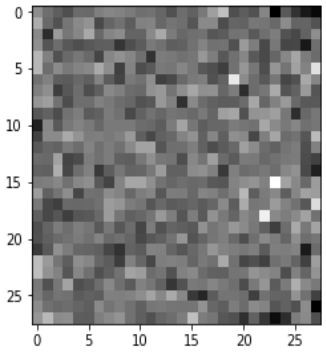
C = 0.1



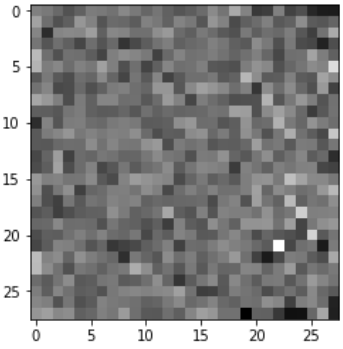
C = 1



C = 10



C = 100



Observation

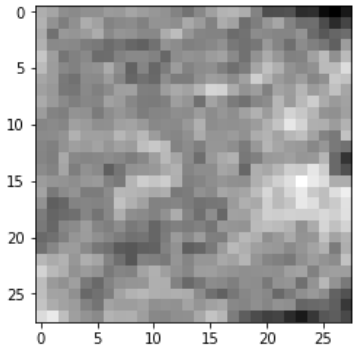
As we can from the above plots as the C value increases the weights tend to distribute and in case of 0.01 the solution is sparse which suggests only few elements present in W are taking part in making decision whereas for C >= 1 the weights are more distributed and regularized .

L2 loss

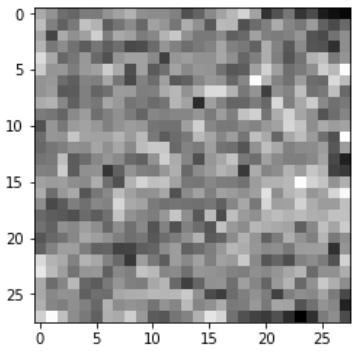
Accuracy on test data with L2 penalty: 96.26%

[Parameters - tolerance = 0.01, C = 0.01]

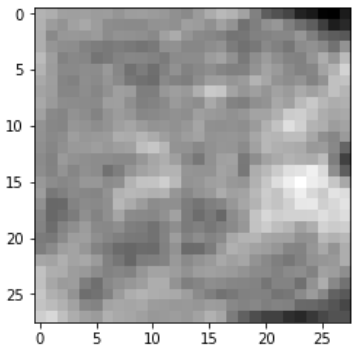
C= 0.01



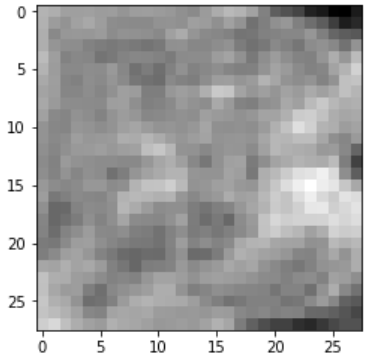
C = 0.1



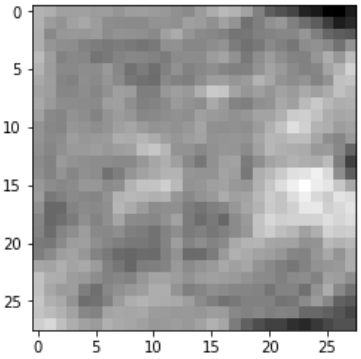
C = 1



C = 10



C = 100



Observation

From the above plots we can infer that as the value of C increases the weights tend to learn a pattern which becomes prominent for values greater than C>= 1. For higher values of C we do not see any prominent change indicating the elements of weights tends to saturate faster . But, in the case L2 loss we see that the classifier is learning a pattern to classify weights whereas the pattern is not so prominent in L1 loss.