

MSCI 641

Assignment 4

Vyas Anirudh Akundy
Student ID: 20765080

1 Classification Accuracy:

The drop out was varied between 0.3, 0.6 and 0.9 and the L2 regularization value was varied between 0.05 and 0.005

Activation Function	Dropout	L2 Reg	Test Accuracy(%)
Sigmoid	-	-	79.28
	0.3	-	79.18
	0.6	-	78.75
	0.9	-	77.74
	-	0.05	78.69
	-	0.005	78.99
	0.3	0.05	79.00
	0.3	0.005	79.40
	0.6	0.05	78.64
	0.6	0.005	78.82
	0.9	0.05	77.50
	0.9	0.005	77.70
ReLU	-	-	78.50
	0.3	-	79.10
	0.6	-	78.21
	0.9	-	77.40
	-	0.05	76.00
	-	0.005	77.16
	0.3	0.05	73.90
	0.3	0.005	76.71
	0.6	0.05	75.80
	0.6	0.005	76.90
	0.9	0.05	75.80
	0.9	0.005	75.20
tanh	-	-	78.50
	0.3	-	78.30
	0.6	-	77.80
	0.9	-	77.56
	-	0.05	77.30
	-	0.005	77.50
	0.3	0.05	75.80
	0.3	0.005	76.60
	0.6	0.05	74.96
	0.6	0.005	75.80
	0.9	0.05	74.50
	0.9	0.005	76.80

2 Reasoning:

2.1 Effect of Activation Functions:

- From the above table we can see that on average, the test accuracy performance for the different cases with different activation functions is,
 - Sigmoid = 78.72%
 - ReLU = 76.72%
 - tanh = 76.78%
- From these values we can clearly see that the Sigmoid activation function is performing slightly better than the other two.
- The reason maybe due to the fact that ReLU better works for deep neural networks to avoid the vanishing gradient problem.
- However the current network architecture has only 2 layers thus Sigmoid seems like a better choice.

2.2 Effect of L2 Regularization:

- Overall, from the data, we can see that introducing L2 regularization has dropped the test accuracy by 1 – 2%, however between the two values, 0.005 seems to perform slightly better.
- Regularization may not improve the accuracy, but it will however improve the generalization of the model

2.3 Effect of Dropout:

- Introducing dropout alone seems to slightly reduce the accuracy in few cases.
- My intuition is that since dropout causes random nodes to be dropped, and these nodes being words in the sentence, while predicting the sentiment, dropping certain words might negatively effect the output.
- Example, important words like "good", "like", "worst" might get dropped resulting in a wrong classification.
- However, for a combination of dropout and L2 regularization, there seems to be a better performance in accuracy than when applied alone.