Neural Networks

Vijay

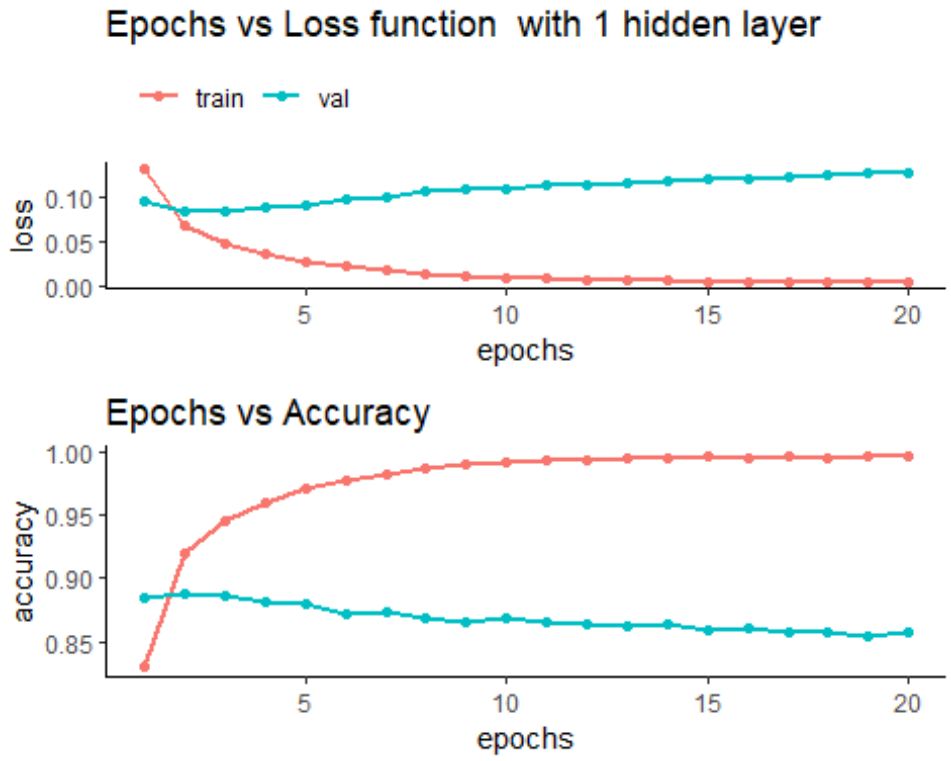
2/3/2020

IMDB dataset, a set of 50,000 highly polarized reviews from the Internet Movie Database. They're split into 25,000 reviews for training and 25,000 reviews for testing, each set consisting of 50% negative and 50% positive reviews.

```
#importing imdb movie review dataset with top 10000 words
imdb <- dataset_imdb(num_words = 10000)
#spliting the data into train and test
c(c(train_data, train_labels), c(test_data, test_labels)) %<-% imdb
```

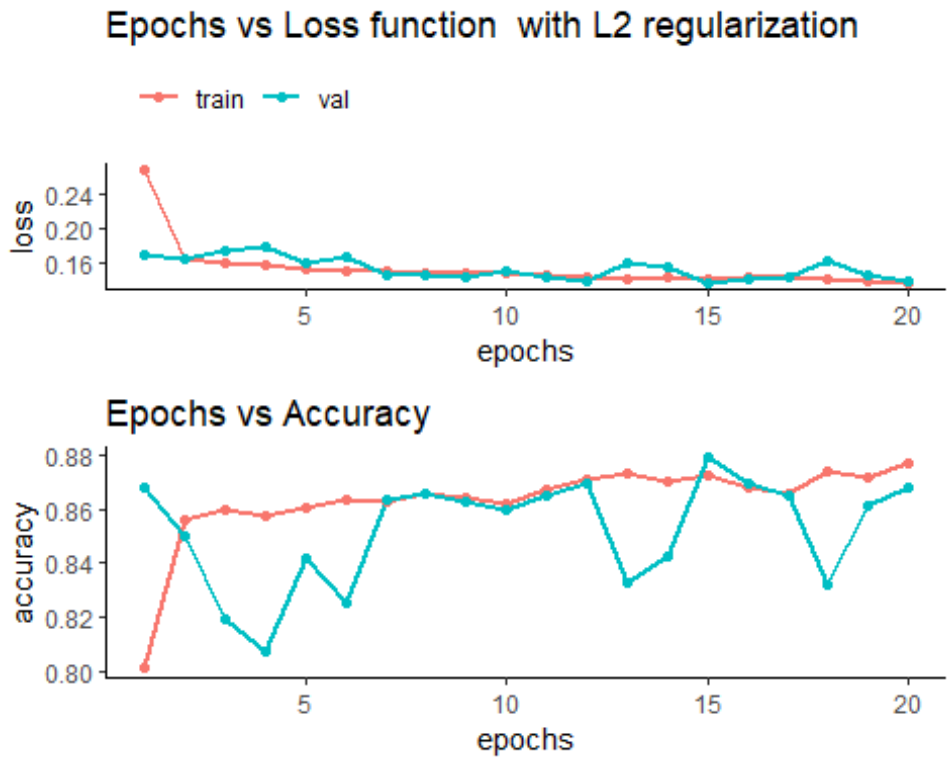
Post partition of data, it is transformed into the datatype which is suitable to the tensors.

Building a neural network model with 32 units, 20 epochs, and a batch size of 250



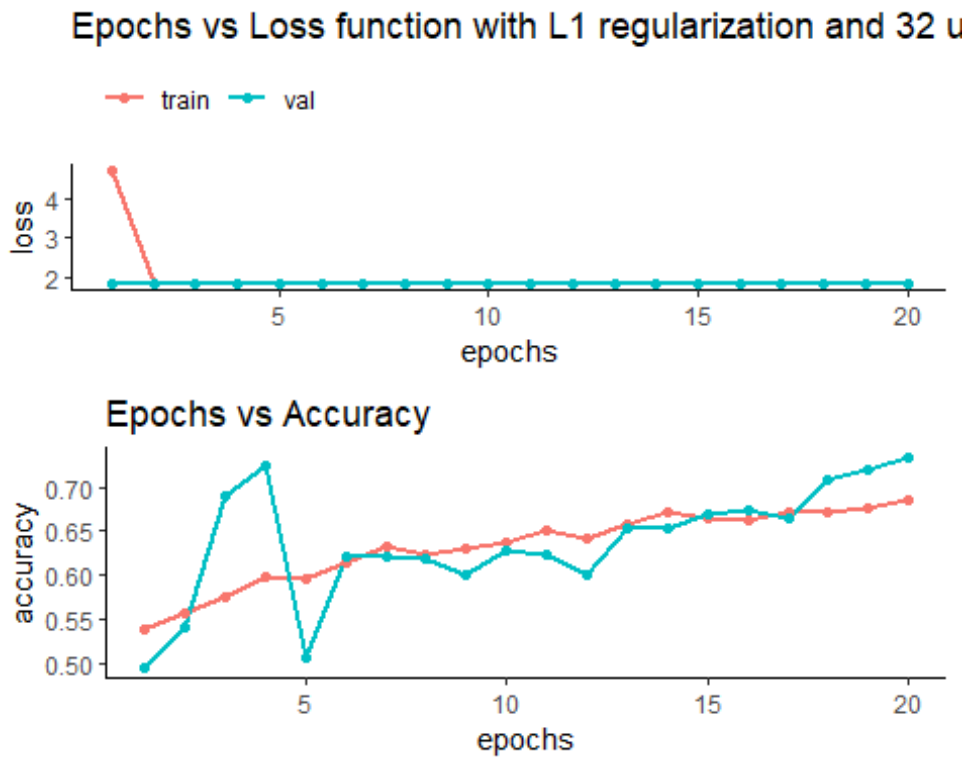
It is clear from the plot that the model accuracy is more in the training data set than the validation data set. Hence the model is "overfit".
By using the Brute force method for units and batch_size, we observed values of 64 and 512 resp. tend to overfit. Hence the model needs to be tuned further.

Examining the neural network model with regularization ("L2 form") and dropout (50%) methods.



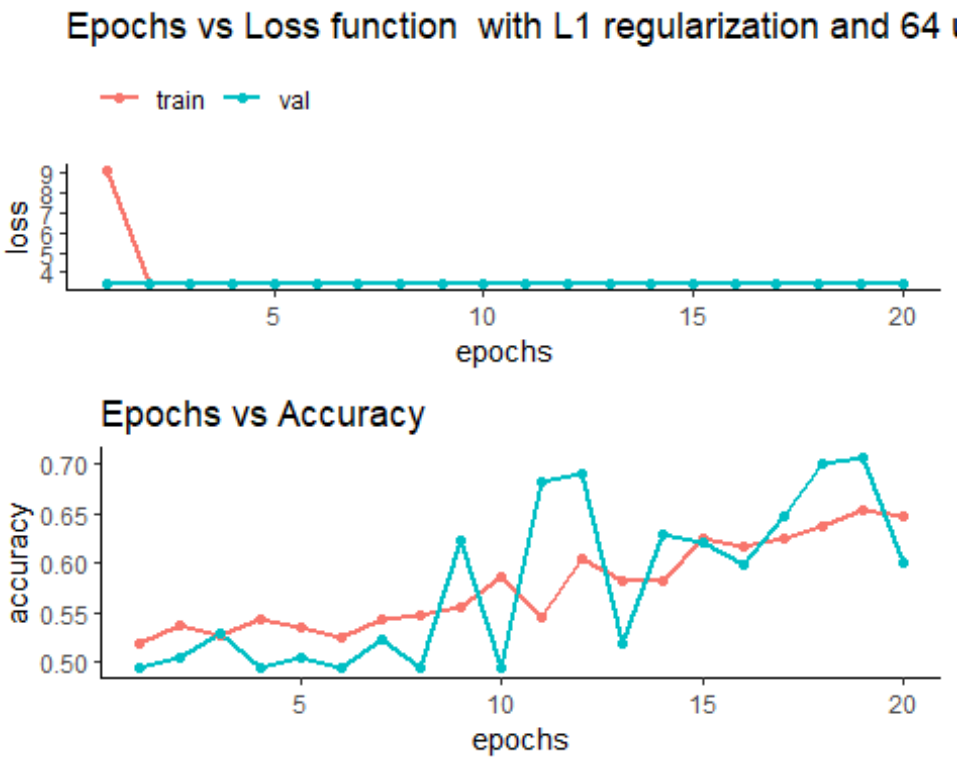
It depicts from the above plot is L2 form of regularization tends to "overfit".
It is observed that the model tends to overfit with 64 units and 250 batch size

Building a neural network model with regularization ("L1 form") and dropout methods.

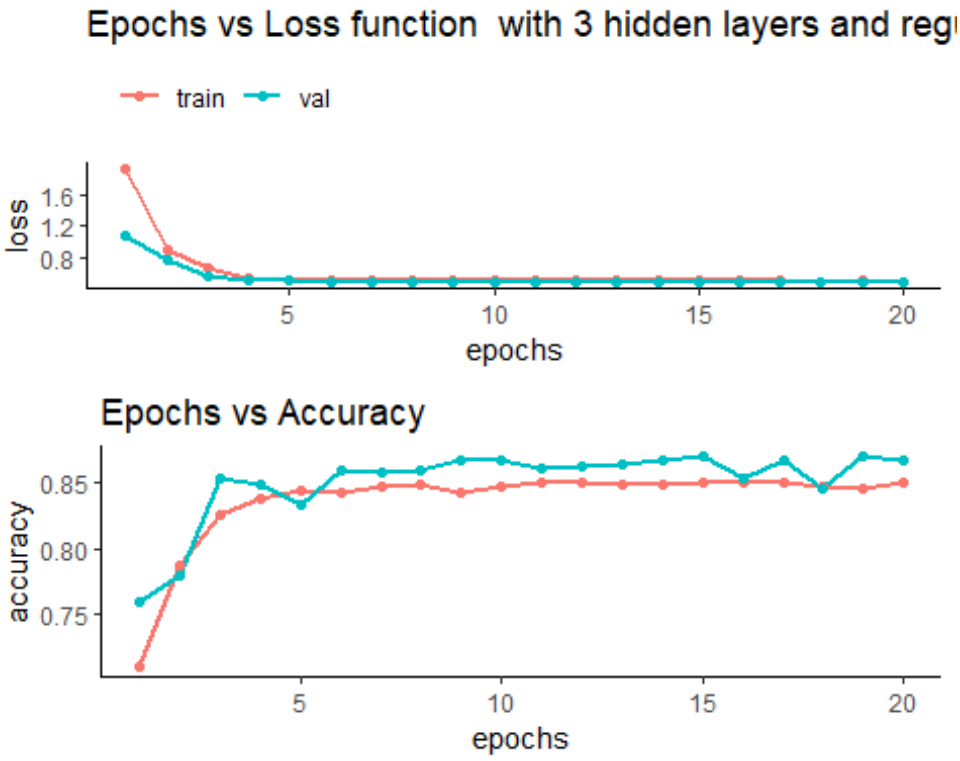


From the above plot, the model seems to perform fairly well.

Using the weights and functions, fine-tuning the model with different inputs.

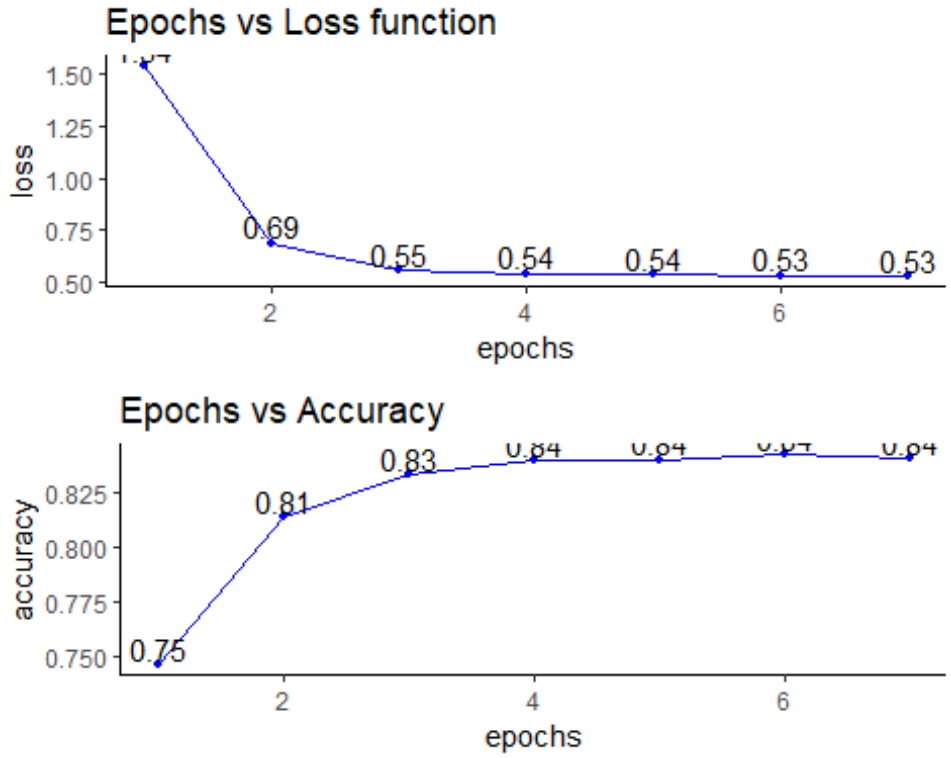


Adding more hidden layers to the model in order to verify the model to perform well



Observed the units with 32 but 64 units tend to perform well. Hence choosing the above m odel as my best model.
Choosing epoch value as 7 since its validation started peaking at 7th epoch

Evaluating the model with test data considering the tuned parameters i.e. 64 units at every 3 hidden layers., epoch as 7 and batch size as 250.




```
Result:

loss
0.5095245
accuracy
0.8638
```