|  |
| --- |
| **Assignment** : Advance Machine Learning Assignment 1   * **Author** : Khushboo Yadav * **Data** : IMDB dataset * **Goal** : To explore and extend original Neural Network model * **Document**: Summary report |

**Introduction:**

***Classifying movie reviews: a binary classification example***

Two-class classification, or binary classification, may be the most widely applied kind of machine-learning problem.

The Goal is to design a model to classify movie reviews as positive or negative, based on the text content of the reviews.

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.

**Requirement:**

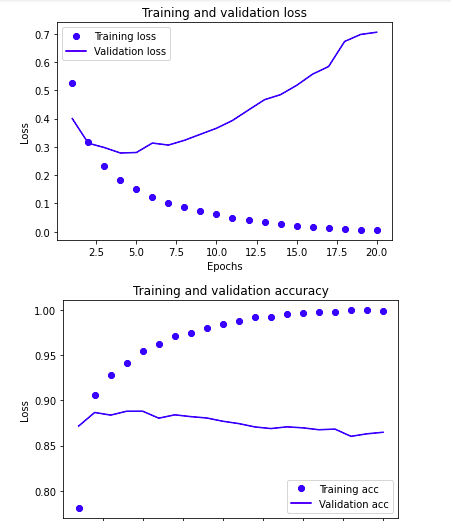
To modify an existing neural network model to improve performance

Explain how different approaches affect the performance of the model

**Problem**: Overfitting of the Original Model

 The original model performs better on the training data but not on data it has never seen before.

This situation is called as Overfitting. It happens when we are we are over-optimizing on the training data, and we ended up learning representations that are specific to the training data and do not generalize to data outside of the training set.



**Solution :**

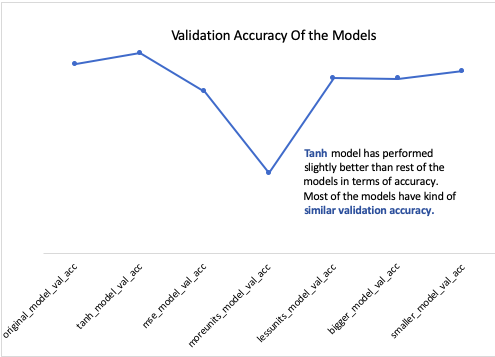
In order to improve the performance of the model we have to follow 4 different approaches.

Below is the pictorial representation of the 4 different approaches to be followed on the original model.

After Building all these different types of models and evaluation of the results. I was able to identify few methods which performed slightly better than original Model. However, there was not significant improvement.

Let’s look at the charts:

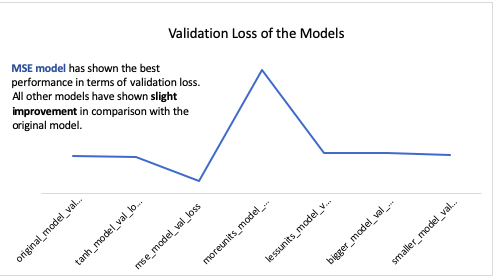
Below line chart shows the average Validation Accuracy of the retrained model for all types of Model.



**Observations**:

1. Tanh model has performed slightly better than rest of the models in comparison with the original model in terms of accuracy.
2. Most of the models have similar validation accuracy in comparison with original model.

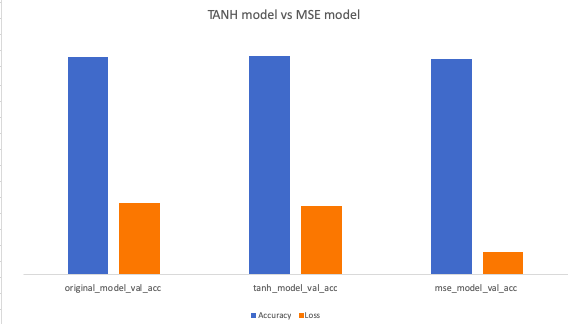
Below line chart shows the average Validation loss of the retrained model for all types of Model.



**Observations**:

1. As per the data, MSE model has outperformed by reducing the validation loss to 9% from 29%.
2. All other models have shown slight improvement; however, the improvement is not that significant enough.

Now, Let’s compare the tanh model and MSE model with original model.



* Tanh Model performed well in accuracy and second best in Validation loss.
* MSE model performed second best and best in Validation loss
* In comparison to Both models with original model, MSE model has shown significant improvement in Validation loss.
* Therefore, MSE is the model with overall best performance.

**Should we declare MSE model as winner …?**

No, we also need to understand why MSE Model has shown significant improvement?

How does it work? Is it good for binary classification model?

Mean squared error is calculated as the **average of the squared differences between the predicted and actual values**. The result is always positive regardless of the sign of the predicted and actual values and **a perfect value is 0.0.** The squaring means that larger mistakes result in more error than smaller mistakes, meaning that the **model is punished for making larger mistakes.**

* Is it good for regression model?

MSE loss function is typically used for regression model, Here IMDB dataset is a classification problem.

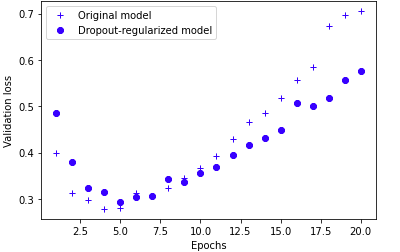
* Comparison with Binary Crossentropy :

if **MSE** is **used**, **it will** return 0 as a loss value, whereas the **binary** cross-entropy **will** return some "tangible" value.

Binary cross entropy intended for use with binary classification where the target values are in the set {0, 1}.

Dropout regularization shows a **real improvement** in the performance of the model.

Problem of overfitting has been handled properly as we can see that difference between training loss and validation loss has been reduced.



**Conclusion: -**

The Architecture of the original model network is better than the new models.

However, we can improve its performance to some level.

With the help of regularization methods i.e. weight regularization, dropout methods, we can overcome overfitting and deliver a better model. The Dropout Model had some positive changes in the results.