ASSIGNMENT 1

NEURAL NETWORKS

SUMMARY REPORT

Combination of hyperparameters to improve test and validation precision.I started by importing every module needed to create and run a model. The IMDB dataset was then loaded.The dataset is split into two sections: a training section and a testing section. Every piece hasIt has 25,000 reviews.The data is then vectorized because we cannot input integers into neural networks after that.Additional training set is broken into partial training and validation sets in two parts.

**MODEL:**

Model construction uses 16 hidden units, the ReLU activation function, the rmsprop optimizer, and the binary cross entropy loss function for the first three models. Each model's number of hidden layers has been modified. The accuracy for training, validation, and testing is One hidden layer 0.9450 0.8894 0.8824 Two hidden layers 0.9983 0.8696 0.8582 Three hidden layers 0.9984 0.8440 0.8326.Given the small number of data samples in the IMDB dataset, we may say that one hidden layer outperforms the other two layers. Overfitting brought on by multiple layers causes a decline in test and validation accuracy.

The next two models are constructed using a single hidden layer, the ReLU activation function, the rmsprop optimizer, and the binary cross entropy loss function. The hidden units are now 32 and 64 in number. The outcomes are 16 hidden units 0.9450 0.8894 0.8824 32 hidden units 0.9925 0.8731 0.8615 64 hidden units 0.9906 0.8698 0.8591.Increasing the hidden units has the same overfitting issue as increasing the hidden layers. The optimal combination to employ is 16 hidden units because it produces the highest accuracy.Change the loss function to mse and compare it to binary cross entropy to check the accuracy now. One hidden layer, 16 hidden units, the ReLU activation function, and the rmsprop optimizer make up these models. Combination Training accuracy Validation accuracy Test accuracy is given as Binary cross entropy 0.9450 0.8894 0.8824 MSE 0.9830 0.8782 0.8701

**DROPOUT:**

I applied the Drop out (0.5) to the following model. I utilized 16 hidden units, one hidden layer, the rmsprop optimizer, the binary cross entropy loss function, and the relu activation function in this model. Combination Training accuracy, Validation accuracy and Test accuracy are written as relu 0.9450 0.8894 0.8824 tanh 0.9925 0.8737 0.8637

Using dropouts causes accuracy to decline. This may be the case since the model is not overfitting and the use of dropout is unfavorable in such circumstances.

**REGULARIZERS:**

The following two models are constructed using regularizers: one hidden layer, sixteen hidden units, the rmsprop optimizer, the binary cross entropy loss function, and the relu activation function. Performance of L1 and L2 regularizers are compared.

Combination Training accuracy ,Validation accuracy and Test accuracy are 0.9450 0.8894 0.8824 L1 regularizer 0.9993 0.8759 0.8667 L2 regularizer 0.9925 0.8798 0.8706.It is clear from the above observation that regularizers are not required for this data because their use does not increase the accuracy.

**OPTIMIZER:**

contrasting rmsprop with adam optimizer. A binary cross entropy loss function, 16 hidden units, a relu activation function, and a hidden layer are all used in the construction of the model.Combination Training accuracy, Validation accuracy and Test accuracy are rmsprop 0.9450 0.8894 0.8824 Adam 0.9816 0.8833 0.8721.Rmsprophas performed better than Adam for the givenIMDBdataset

In summary, a number of variables influence how well a model performs in a neural network. The first is the size of the sample of data. In order to improve accuracy, we require additional data samples.

Now, one layer with 16 hidden units outperforms other hidden layers in this case with 25 000 samples. Given that the ReLU function gives the model a benefit.

I discovered that we don't actually need to utilize it until it is necessary when I noticed that the accuracy was declining when I tried using various optimizers and regularizers.

The model might perform better in the case that we train with numerous layers, utilize regularizers, and dropouts while working with a large amount of data.

Lastly, for the ideal configuration of hyper parameters for the provided dataset includes one hidden layer, 16 hidden units, the Binary cross entropy loss function, the RELU activation function, and the RMSP optimizer.