

Dr. Tony Diana  
DATA 602 Introduction to Machine Learning  
Homework I Week 11

---

**Exercise 1.**

1. Install preprocessing with Keras
2. Load all the necessary libraries
3. Load tensorflow.keras.models
4. Load the dataset using from keras.datasets import imdb
5. Limit the number of words to 10,000 in the training and testing datasets
6. Print the number of unique categories for the target variable
7. Print the average review length
8. Print the standard deviation
9. Vectorize sequences and dimension with a size of 10,000
10. Make sure the test and train variables represent 10,000 observations
11. Import Sequential, Flatten, Dropout, and Dense from tensorflow.keras.models
12. Use the following parameters with an 'adam' optimizer and 'binary\_crossentropy' loss function:

Model: "sequential"

Layer (type)	Output Shape	Param #
=====	=====	=====
dense (Dense)	(None, 50)	500050
dropout (Dropout)	(None, 50)	0
dense_1 (Dense)	(None, 1024)	52224
dropout_1 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 2)	2050
dense_3 (Dense)	(None, 1)	3
=====	=====	=====
Total params: 554,327		
Trainable params: 554,327		
Non-trainable params: 0		

---

13. Compile, train, and evaluate the model with epochs=20, batch\_size=512, validation\_data=(test\_x,test\_y)
14. Provide the loss and accuracy for 20 epochs

## Exercise 2.

1. from keras.preprocessing.text import Tokenizer
2. Use keras to build a 'Sequential' model
3. Use the following model:  
  
epochs = 20  
maxlen = 100  
embedding\_dim = 50  
num\_filters = 64  
kernel\_size = 5  
batch\_size = 32
4. Use the 'yelp\_labelled.txt'
5. Use 'sentence' as feature and 'label' as target
6. Tokenizer with 5,000 words
7. Pad the X\_train and X\_test values
8. Use a 'Sequential()' model with

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 50)	84900
conv1d (Conv1D)	(None, 96, 64)	16064
global_max_pooling1d (Global	(None, 64)	0
dense_4 (Dense)	(None, 10)	650
dense_5 (Dense)	(None, 1)	11
Total params: 101,625		
Trainable params: 101,625		
Non-trainable params: 0		

9. Determine the loss, accuracy of the training and testing model