In [1]:
```
# import libraries
import sklearn
from sklearn import model_selection
from sklearn.model_selection import train_test_split
from sklearn import preprocessing
from sklearn.preprocessing import OneHotEncoder
import tensorflow as tf
import keras
import re
import tensorflow as ft
from tensorflow import keras
from keras import preprocessing
from tensorflow.keras.callbacks import EarlyStopping
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential
from tensorflow.keras.models import load_model
from tensorflow.keras.layers import Dense, Embedding
import sys
import os
import pandas as pd
import numpy as np
import gzip
import re
import seaborn as sns
import nltk
import re
from nltk.corpus import stopwords
from nltk import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
nltk.download ('stopwords')
nltk.download ('punkt')
nltk.download ('wordnet')
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\Todd\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Todd\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\Todd\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
```

Out[1]:
```
True
```

In [2]:
```
#B - Exploratory data analysis

#1
# read in data
dfYelp = pd.read_excel("C:\\Users\\Todd\\Desktop\\Desktop\\WGU\\Advanced Data Analytics\\Ye
dfYelp.describe()
```

Out[2]:

	label
count	1000.00000
mean	0.50000
std	0.50025
min	0.00000