from imblearn.over\_sampling import SMOTE

counter = Counter(y\_train)

print('Before',counter)

# oversampling the train dataset using SMOTE

smt = SMOTE()

#X\_train, y\_train = smt.fit\_resample(X\_train, y\_train)

X\_train\_sm, y\_train\_sm = smt.fit\_resample(X\_train, y\_train)

counter = Counter(y\_train\_sm)

print('After',counter)

def remove\_punctuation(text):

import string

translator = str.maketrans('', '', string.punctuation)

return text.translate(translator)

<https://gist.github.com/slowkow/7a7f61f495e3dbb7e3d767f97bd7304b>