5.10.2024

完成限制生成的对抗样本扰动在20%以下，做了如下修改

1. d:\\TextAttack\textattack\attack.py 在后面添加了这样一段代码

class MyMaxModificationRateConstraint(Constraint):  
 def \_\_init\_\_(self, max\_rate, min\_threshold=1):  
 assert isinstance(max\_rate, float), "`max\_rate` must be a float."  
 assert max\_rate >= 0 and max\_rate <= 1, "`max\_rate` must be between 0 and 1."  
 assert isinstance(min\_threshold, int), "`min\_threshold` must be an int."  
  
 self.max\_rate = max\_rate  
 self.min\_threshold = min\_threshold  
 self.compare\_against\_original = True # 添加 compare\_against\_original 属性  
  
 def \_check\_constraint(self, current\_text, transformed\_text):  
 threshold = max(  
 math.ceil(current\_text.num\_words \* self.max\_rate), self.min\_threshold  
 )  
 if len(current\_text.attack\_attrs["modified\_indices"]) >= threshold:  
 return set()  
 else:  
 return set(range(len(current\_text.words)))  
  
 def extra\_repr\_keys(self):  
 return ["max\_rate", "min\_threshold"]

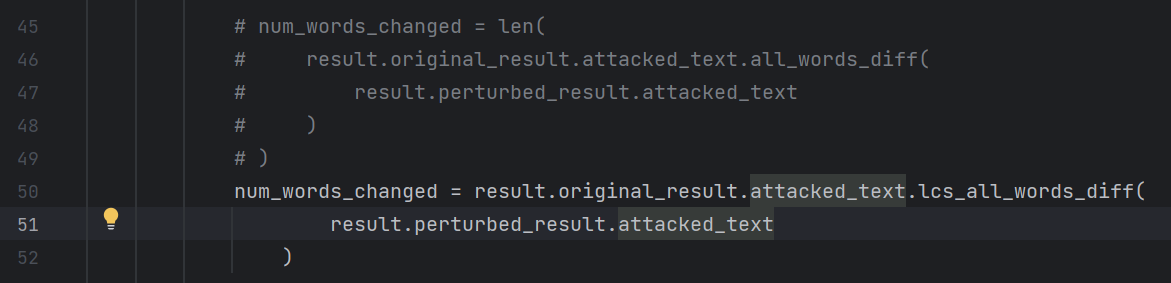
并将第124行 从self.constraints = []修改为self.constraints = [MyMaxModificationRateConstraint(0.2)]

5.16

关于扰动率计算的修改

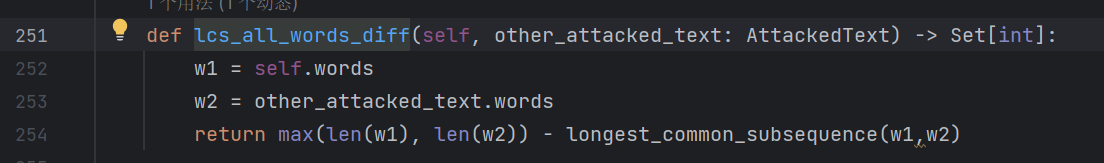
1. D:/TextAttack/textattack/metrics/attack\_metrics/words\_perturbed.py

将原始45-49行注释掉，添加了50-52行



1. D:/TextAttack/textattack/shared/attacked\_text.py

添加了如下几行代码



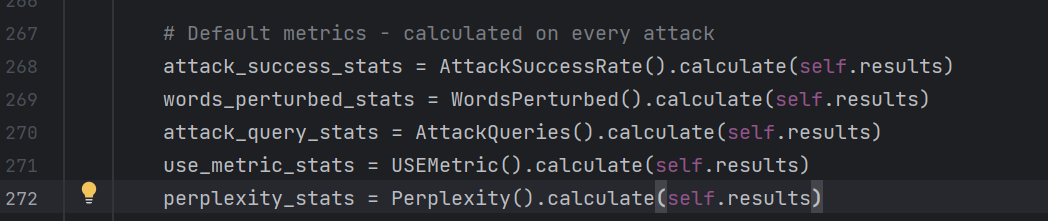
1. 要保留这个D:\TextAttack\testdata\dOpenHowNet.py文件

5.20

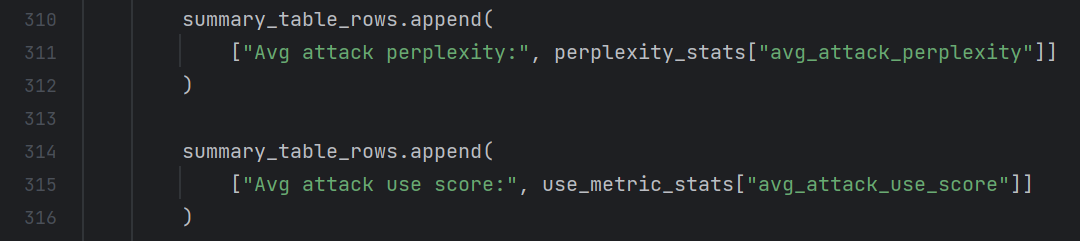
为了输出指标use语义相似度和perplexity模型困惑度，作了如下修改

1. D:\TextAttack\textattack\loggers\attack\_log\_manager.py

添加了271和272行



添加了这两段



1. 

修改了本地路径

