

归一化后挑出分数差较大的

Human score极高(0.997) automated metrics 低(0.59)

- Ref: an enjoyable , if occasionally flawed , experiment
- Out: an nice , if occasionally faulty , experiment

可能的原因：句子短，改动大？

automated metrics分布

0.91351008 0.56888499 0.55582209 0.92062334 0.90341671 0.55582209 0.52329758 0.31975587 0.63155219 0.83714181 5.75E-05 0.37869704

Human score高(0.838) automated metrics低(0.483)

- Ref: An animal is in water
- Out: An critters is in water

原因：没理解同义词

automated metrics分布

0.46723655 0.63461778 0.64547081 0.4872553 0.4430869 0.64547081 0.36920242 0.23593789 0.62131752 0.73512246 9.09E-05 0.51320839

Human score低(0.79) automated metrics高(0.93)

- Ref:we always have a good time whenever we hit the luxor ! before , when it was still called ' ra ' it was crowded , it was fun , and there was always someone new to meet n nmy girls and i unknowingly planned our trip to las vegas for memorial day weekend i say unknowingly because i know better than to go to las vegas during a holiday weekend too many people \ (i just do n't do big crowds \) , too many drunks , and clubs become too expensive however , it was n't my birthday and that weekend worked so we went n nwe hit lax the last night of our stay in vegas \ (sunday night \) we got scammed by some dude selling club passes on the strip \ (in front of caesar 's , do n't give him your money \) so we knew it would a rough night of clubbing ahead of us we decided to just go for it and headed to the luxor we were n't on a list , we did n't have passes , we were doing it ' the regular way ' which i have n't done since i first started clubbing \ (about 10 years ago \) we expected the cover to be insane , especially when we discovered that ginuwine was performing \ (do n't judge me , i still love ginuwine \) we were told that the cover was 20 , but we could expect to pay more by the time we got to the door totally understandable , this is vegas , the goal is to make money n nwe got to the door in record time \ (less than an hour of waiting in line , score ! \) , and we only paid 30 to get in i look at this as a bargain because it was a holiday weekend , other clubs were charging much more and there was a performance that night n nthe only real problem we had was that the club was so packed it was really hard to finish my drink without spilling it on myself \ (thanks to the thousands of people milling back and forth bumping in to me \) after ginuwine performed a lot of people left so it was easier to dance easier , not easy it was still super packed , no real dancing could be done and getting off the dance floor was very tricky n nover all , we had a great time the luxor never disappoints me it 's always an adventure when we roll through the luxor
- Out:we always have a okay time whenever we hit the luxor ! before , when it was still called ' ra ' it was crowded , it was fun , and there was always someone new to meet n nmy girls and i unknowingly planned our trip to las vegas for memorial day weekend i say unknowingly because i know better than to go to las vegas during a holiday weekend too many people \ (i just do n't do big crowds \) , too many drunks , and clubs become too expensive however , it was n't my birthday and that weekend worked so we went n nwe hit lax the last night of our stay in vegas \ (sunday night \) we got scammed by some dude selling club passes on the strip \ (in front of caesar 's , do n't give him your money \) so we knew it would a rough night of clubbing ahead of us we decided to just go for it and headed to the luxor we were n't on a list , we did n't have passes , we were doing it ' the regular way ' which i have n't done since i first started clubbing \ (about 10 years ago \) we expected the cover to be insane , especially when we discovered that ginuwine was performing \ (do n't judge me , i still love ginuwine \) we were told that the cover was 20 , but we could expect to pay more by the time we got to the door totally understandable , this is vegas , the goal is to make money n nwe got to the door in record time \ (less than an hour of waiting in line , score ! \) , and we only paid 30 to get in i look at this as a bargain because it was a holiday weekend , other clubs were charging much more and there was a performance that night n nthe only real problem we had was that the club was so packed it was really hard to finish my drink without spilling it on myself \ (thanks to the thousands of people milling back and forth bumping in to me \) after ginuwine performed a lot of people left so it was easier to dance easier , not easy it was still super packed , no real dancing could be done and getting off the dance floor was very tricky n nover all , we had a great time the luxor never disappoints me it 's always an adventure when we roll through the luxor

可能原因：句子过长，改动较小

Human score高(0.83) automated metrics极低(0.39)

- Ref: totally messed up my order did not care it was crazy how much they did not care was kind of like hey you drove here so might as well take what we got
- Out: heartily messed up my orders get not sympathy it was hysterical how immensely they gets not caring was kinds of like ooh you driven here however may as also taking what we come

不知道为什么human score这么高？标注问题？

Human score较低(0.36) automated metrics
极低(0)

- Ref: A man is yelling at his son about the broken television behind them .
- Out: A friend is cheering at his threading about the infringed custard behind them .

Human score 不是特别低的原因是因为场景比较类似？只是实体变了？

Human score高(0.717) automated metrics极低(0.26)

- Ref: The building is old .
- Out: The buildings is immemorial

可能原因：机器认为旧和远古的这两个词意思区别较大

automated metrics分布

0.52384508	0	0.28687592	0.54286512	0.50134519	0.28687592	0	0.14055722	0.50050308	0.36432639	0	0
------------	---	------------	------------	------------	------------	---	------------	------------	------------	---	---

Human score低(0.312) automated metrics高(0.74)

- Ref: A woman has her shoe off to adjust her sock .
- Out: A woman has her galoshes off to adjust her sock .

可能原因：生成句子不符合事实

automated metrics分布

0.68793764 0.85571844 0.84106802 0.69938343 0.67520104 0.84106802 0.60997529 0.42233503 0.85198689 0.92470556 0.70570913 0.78919232

Human score低(0.306) automated metrics高(0.767)

Out: mike was our server he was knowledgeable great personality very attentive the food and his service was awesome we will definitely be back and ask for him

Ref: mike was our server he was knowledgeable great personality very attentive the food and his service was funky we will soberly constituted back and ask for him

动词区别很大，求婚和回来找他

0.75769821 0.85021537 0.80484632 0.7626803 0.75360385 0.80484632 0.60672601 0.48593074 0.83173394 0.9805635 0.77057584 0.80622478

Human score低(0.07) automated metrics高(0.733)

Ref: the women are looking at a bouquet of flower

Out: the women are striving at a bouquet of flower

动词的意义区别很大

0.75261843 0.81928803 0.80484632 0.75982818 0.74536965 0.80484632 0.62801435 0.38264045 0.81524056 0.96002013 0.60037014 0.72823419

Human score高(0.84) automated metrics低(0.43)

Ref: The man does not have any gear on .

Out: The man gets not had any accelerating on .

汽油和加速有着同样作用，所以human判断相关？

0.72952459 0.40964401 0.40640755 0.74165558 0.7153678 0.40640755 0.14033692 0.21085699 0.58200712 0.87767283 7.98E-09 2.89E-06