

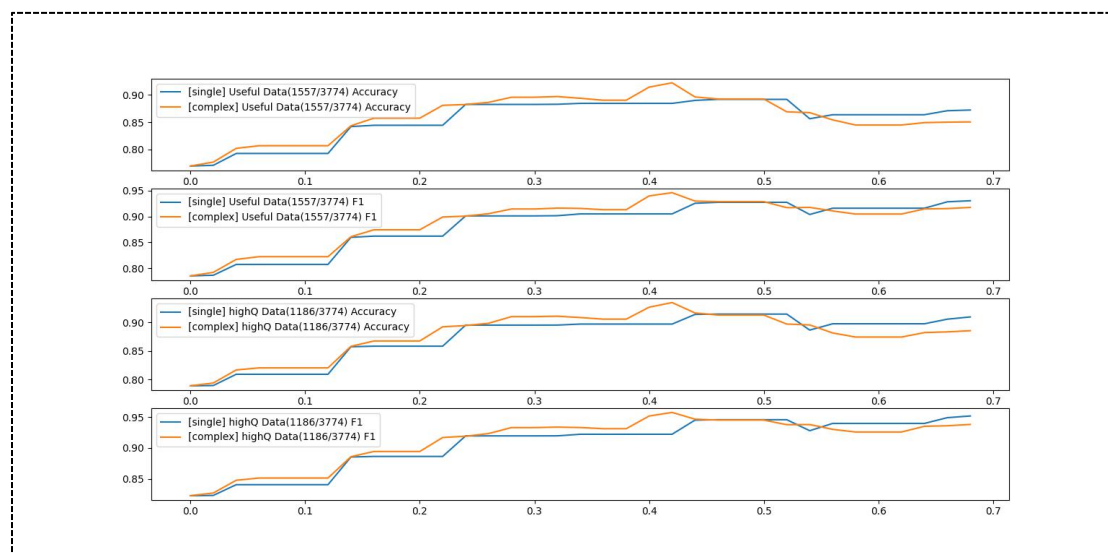
2021.08.05 ~ 2021.08.21 工作进展

1 单阈值 -> 阈值区间

核心算法：之前设定单个距离阈值 `threshold`，判断说对与否。现在改为设定阈值区间`[a, b]`，距离 $< a$ 判定为说对，距离 $\geq b$ 判定为说错， $a \leq \text{距离} < b$ 系统不作判定，交给人工。

区间的确定：将 4k 条测试数据得到的距离正序排序，例如得到`[0, 0, 0.5, …… , 4.0]`。假设人工判定的比例为 30%，即 $4k * 0.3 = 1.2k$ 个距离需要人工判定，系统判为对的比例为 `correct`，则系统判定为错的比例为 $0.7 - \text{correct}$ ，每次按照设定的 `step` 值增大 `correct`，求解使准确率（F1 值）最高的 `correct`，然后映射为`[a, b]`。

2.1 人工判定 30%，step = 0.02



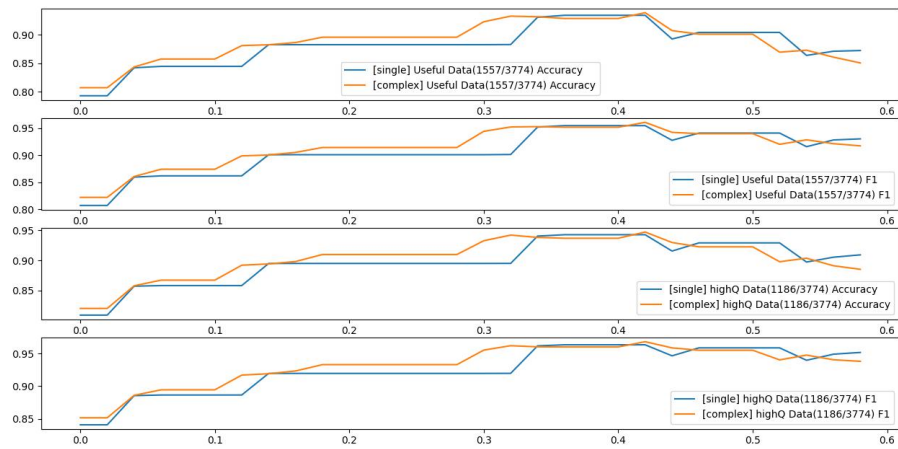
```
-----All Data Best Performance-----
[single] threshold: [1.50, 2.50] , accuracy: 89.17%
[complex] threshold: [1.00, 2.25] , accuracy: 92.18%

[single] threshold: [2.00, 4.00] , F1: 93.05%
[complex] threshold: [1.00, 2.25] , F1: 94.61%

-----HighQ Data Best Performance-----
[single] threshold: [1.50, 2.50] , accuracy: 91.43%
[complex] threshold: [1.00, 2.25] , accuracy: 93.44%

[single] threshold: [2.00, 4.00] , F1: 95.17%
[complex] threshold: [1.00, 2.25] , F1: 95.74%
```

2.2 人工判定 40%， step = 0.02



```
-----All Data Best Performance-----  
[single] threshold: [1.00, 2.50] , accuracy: 93.39%  
[complex] threshold: [1.00, 2.60] , accuracy: 93.85%  
  
[single] threshold: [1.00, 2.50] , F1: 95.50%  
[complex] threshold: [1.00, 2.60] , F1: 96.11%  
  
-----HighQ Data Best Performance-----  
[single] threshold: [1.00, 2.50] , accuracy: 94.28%  
[complex] threshold: [1.00, 2.60] , accuracy: 94.72%  
  
[single] threshold: [1.00, 2.50] , F1: 96.33%  
[complex] threshold: [1.00, 2.60] , F1: 96.81%
```