**بسم الله الرحمن الرحیم و توکلت علی الله**

**گزارش دقت سیستم تخصیص عناصر صحنه در حالات مختلف**

معیار ارزیابی: بر اساس عنصر صحنه‌های ~~یادگرفته شده توسط CRF~~ (نه زیرا در این کار هدف اندازه گیری دقت سیستم تخصیص است نه تشخیص عنصر صحنه)، لیبل خورده در کورپس

۱- در حالتی که تنها بر اساس عنصر صحنه‌های لیبل خورده در کورپس، هر عنصر در بخش مربوطه از مدل صحنه قرار میگیرد.

در این حالت ضمایر چون عنصر صحنه ندارند بدرستی تخصیص داده نمیشوند.لذا همه آنها در FN قرار میگیرند و از ریکال می کاهند.

لازم به ذکر است در این حالت 497 تعداد همه کلمات در همه عناصر صحنه (۱۲ عنصر) در کل مدل صحنه‌ها در مدل صحنه طلایی است.

total (not NO and JUNK) word nums: 497

weighted accuracies: 86.27910453828845

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RoleEmoent evaluation:  TP:0 FN:0 FP:0 TN:1203 | RoleIntent evaluation:  TP:70 FN:0 FP:0 TN:1133  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 1203 roleIntent Number: 70 | RoleState evaluation:  TP:56 FN:0 FP:0 TN:1147  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 1203 roleState Number: 56 | RoleAction evaluation:  TP:72 FN:0 FP:0 TN:1131  precision:100.0 recall: 100.0 F1\_score: 72.28915662650601  total number: 1203 roleAction Number: 72 | Role evaluation:  TP:90 FN:69 FP:0 TN:1044  precision:100.0 recall: 56.60377358490566 F1\_score: 72.28915662650601  total number: 1203 role Number: 159 |
|  |  | DynamicObjectStates evaluation:  TP:5 FN:0 FP:0 TN:1198  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 1203 dynmic\_obejcts\_state Number: 5 | DynamicObjectActions evaluation:  TP:16 FN:0 FP:0 TN:1187  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 1203 dynamic\_obejct\_Actions Number: 16 | DynamicObjects evaluation:  TP:12 FN:1 FP:0 TN:1190  precision:100.0 recall: 92.3076923076923 F1\_score: 95.99999999999999  total number: 1203 dynamic\_obejcts Number: 13 |
|  |  |  | StaticObjectStates evaluation:  TP:10 FN:0 FP:0 TN:1193  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 1203 static\_obejct\_states Number: 10 | StaticObjects evaluation:  TP:23 FN:3 FP:0 TN:1177  precision:100.0 recall: 88.46153846153845 F1\_score: 93.87755102040816  total number: 1203 static\_obejcts Number: 26 |
|  |  |  | AlternativeTimes evaluation:  TP:10 FN:0 FP:0 TN:1193  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 1203 alternative Times Number: 10 | AlternativeLocations evaluation:  TP:56 FN:4 FP:0 TN:1143  precision:100.0 recall: 93.33333333333333 F1\_score: 96.55172413793103  total number: 1203 alternative Locations Number: 60 |

۲- حالتی که مرجع ضمیرها را بصورت دستی مشخص کرده ایم. میزان افزایش دقت سیستم در این مرحله، اهمبت تعیین مرجع ضمیر را مشخص می‌کند.

در این حالت همه پرسیژن و ریکال و اف یک و همه چی ۱۰۰ میشود ☺ بقیه مثل همین رول هستند فقط تعدادشان متفاوت است.

Role evaluation:

TP:159 FN:0 FP:0 TN:1044

precision:100.0 recall: 100.0 F1\_score: 100.0

total number: 1203 role Number: 159

total (not NO and JUNK) word nums: 497

weighted accuracies: 100.0

**گزارش دقت سیستم تشخیص کلمات غیر تکراری در یک مدل صحنه**

در این سیستم همه کلماتی که در همه جملات یک مدل صحنه به یک عنصر صحنه اشاره دارند، اگر بیش از یک کلمه باشند به عنوان عنصر تکراری شناخته شده و اگر تنها یک کلمه به آن عنصر صحنه اشاره کند، به عنوان کلمه غیر تکراری شناخته میشود.

۱- در حالتی مشابه حالت ۱ بالا (در حالتی که تنها بر اساس عنصر صحنه‌های لیبل خورده در کورپس، هر عنصر در بخش مربوطه از مدل صحنه قرار میگیرد.) دقت ها بشرح زیر است. در این حالت ضمایر چون عنصر صحنه ندارند بدرستی تخصیص داده نمیشوند و در تکراری و غیر تکراری بودن عناصر اختلال ایجاد میکنند و برخی از کلمات غیر تکراری را تکراری تشخیص میدهند و لذا همه در FP قرار میگرند و از پرسیژن می کاهند.)

لازم به ذکر است در این حالت 342 تعداد همه کلمات غیر تکراری در کل مدل صحنه ها برای کل عناصر صحنه (۱۲ عنصر) است.

total (not NO and JUNK) word nums: 342

weighted accuracies: 81.62862271718294

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RoleEmoent evaluation:  TP:0 FN:0 FP:0 TN:0 | RoleIntent evaluation:  TP:70 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 70 roleIntent Number: 70 | RoleState evaluation:  TP:56 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 56 roleState Number: 56 | RoleAction evaluation:  TP:72 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 46.35761589403973  total number: 72 roleAction Number: 72 | Role evaluation:  TP:35 FN:0 FP:81 TN:43  precision:30.17241379310345 recall: 100.0 F1\_score: 46.35761589403973  total number: 159 role Number: 35 |
|  |  | DynamicObjectStates evaluation:  TP:5 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 5 dynmic\_obejcts\_state Number: 5 | DynamicObjectActions evaluation:  TP:16 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 16 dynamic\_obejct\_Actions Number: 16 | DynamicObjects evaluation:  TP:7 FN:0 FP:2 TN:4  precision:77.77777777777779 recall: 100.0 F1\_score: 87.50000000000001  total number: 13 dynamic\_obejcts Number: 7 |
|  |  |  | StaticObjectStates evaluation:  TP:10 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 10 static\_obejct\_states Number: 10 | StaticObjects evaluation:  TP:18 FN:0 FP:5 TN:3  precision:78.26086956521739 recall: 100.0 F1\_score: 87.8048780487805  total number: 26 static\_obejcts Number: 18 |
|  |  |  | AlternativeTimes evaluation:  TP:10 FN:0 FP:0 TN:0  precision:100.0 recall: 100.0 F1\_score: 100.0  total number: 10 alternative Times Number: 10 | AlternativeLocations evaluation:  TP:43 FN:0 FP:5 TN:12  precision:89.58333333333334 recall: 100.0 F1\_score: 94.50549450549451  total number: 60 alternative Locations Number: 43 |

۲- حالتی که مرجع ضمیرها را بصورت دستی مشخص کرده ایم. میزان افزایش دقت سیستم در این مرحله، اهمیت تعیین مرجع ضمیر را مشخص می‌کند.

در این حالت همه پرسیژن و ریکال و اف یک و همه چی ۱۰۰ میشود ☺ بقیه مثل همین رول هستند فقط تعدادشان متفاوت است.

Role evaluation:

TP:35 FN:0 FP:0 TN:124

precision:100.0 recall: 100.0 F1\_score: 100.0

total number: 159 role Number: 35

total (not NO and JUNK) word nums: 342

weighted accuracies: 100.0