

Приложение Н.

Результаты для дорожки контекстно-зависимого аннотирования

В этом приложении представлена сводная информация о наилучших результатах по ряду метрик для дорожки контекстно-зависимого аннотирования документов, т.е. максимальный достигнутый результат. Отметим, что для каждой клетки таблицы (то есть категория/метрика) рекордсмен выбирался независимо и поэтому результаты по разным метрикам вполне могли быть получены разными системами.

Информация разбита на три группы соответствующая разным порогам признания аннотации релевантной, использовавшимся при слиянии мнений экспертов. При вычислении оценок использовались два набора матриц релевантности – матрицы, построенные для аннотаций, и матрицы, построенные для документов по которым строились аннотации. Поэтому каждая группа оценок содержит 4 строки (например, *or_and* означает, что используется OR-версия таблицы релевантности для аннотаций и AND версия таблицы релевантности для документов).

Метрика	AAccur	AError	PrecA	PrecA (macro)	PrecD	PrecD (macro)
Порог при слиянии оценок: Relevant-minus						
and_and	.848	.414	.327	.321	.557	.566
and_or	.949	.535	.327	.321	.669	.676
or_and	.768	.311	.598	.603	.557	.566
or_or	.894	.378	.598	.603	.669	.676
Порог при слиянии оценок: Relevant-plus						
and_and	.878	.466	.218	.205	.557	.566
and_or	.956	.598	.218	.205	.669	.676
or_and	.802	.363	.469	.473	.557	.566
or_or	.915	.471	.469	.473	.669	.676
Порог при слиянии оценок: Vital						
and_and	.958	.505	.148	.103	.557	.566
and_or	.983	.635	.148	.103	.669	.676
or_and	.840	.428	.325	.319	.557	.566
or_or	.939	.553	.325	.319	.669	.676

Метрики имеют следующий смысл:

- AAccur (AnnotationAccuracy) - доля релевантных аннотаций (по мнению экспертов), которым соответствуют реально релевантные документы.
Отметим, что дополнение ее до единицы – это доля false positives, то есть случаев, когда пользователь был введен в заблуждение и смотрел лишнюю информацию (нерелевантный документ).
- AError (AnnotationError) – доля нерелевантных аннотаций, которым соответствуют релевантные документы. (То есть случаев, когда пользователь не смотрит документ из-за плохой аннотации)
- PrecA (PrecisionAnnotations) - доля аннотаций признанных релевантными экспертами. То есть точность по таблице релевантности для аннотаций.
- PrecD (PrecisionDocuments) - доля документов, для которых строились аннотации, которые были признаны релевантными при оценке в 2004 году. То есть точность по таблице релевантности для поисковых дорожек.