

П.Г. Полищук: Уверен существует большое число альтернативных решений этой проблемы. Сравнивали ли Вы результаты с другими подходами? Без этого сложно понять хорошие модели или нет.

Н.Ю. Бизюкова: Помимо CRF для извлечения наименований химических соединений (и других объектов, например, биологических) используются и другие методы, такие, как SVM, нейронные сети и методы, основанные на правилах. Иные подходы в литературе встречались реже. Результаты нашей модели сравнивались с остальными: модели, которые обучались на корпусе CHEMDNER, используемом в нашей работе, F1-score лежит в диапазоне между 0,7 и 0,9.

П.Г. Полищук: Почему был выбран именно CRF алгоритм?

Н.Ю. Бизюкова: У алгоритма CRF есть несколько преимуществ для анализа текстов: 1) он позволяет учитывать контекст, что важно для сохранения взаимосвязей между словами в тексте; 2) по сравнению, например, с методами, основанными на правилах, CRF более универсален (несмотря на то, что английский язык более «строгий» в плане строения предложений, нежели русский, в тексте могут использоваться «нестандартные» конструкции). Помимо этого, результаты классификации моделью CRF можно использовать в комбинации с другими методами (например, нейронными сетями).

П.Г. Полищук: Что по Вашему мнению важнее в данных задачах, precision или recall?

Н.Ю. Бизюкова: Precision и recall являются одинаково важными показателями. Идеальной ситуацией является тот случай, когда precision и recall примерно равны и наиболее высоки. К сожалению, добиться их примерного равенства нам не удалось; однако в некоторых случаях при измененных наборах дескрипторов удавалось достигнуть значений precision, близких к единице.

П.Г. Полищук: Как считалась общая точность на слайде 11? Она составляет 0.69, тогда как для индивидуальных задач она только в одном случае ниже этого значения, а в остальных намного выше.

Н.Ю. Бизюкова: Для решения задачи с применением алгоритма для сторонней выборки заранее не аннотированных абстрактов, нами были проставлены метки в четырех столбцах: true positive, true negative, false positive, false negative. Напротив каждого токена выставлялась единица в соответствующий столбец. Затем, для precision, рассчитывалось отношение числа всех единиц в столбце true positive к числу единиц в столбцах true positive и false positive. Для recall - отношение числа всех единиц в столбце true positive к числу единиц в столбцах true positive и false negative. По сравнению с корпусом CHEMDNER, в которые вошли абстракты, посвященные различным темам, в выборку для этого эксперимента вошли абстракты, посвященные только ингибированию ВИЧ. Вполне возможно, что конструкции, используемые в абстрактах по ВИЧ, были плохо представлены в выборке из CHEMDNER.