

Т.И. Маджидов: В качестве меры для выбора 100 мишеней для кластеризации выбрано высокое значение IAP. Но IAP характеризует качество ранжирования вероятности активности, а не сами предсказания (активно/неактивно). Почему не использована метрика, которая характеризует непосредственное предсказание?

Н.С. Ионов: Для отбора белков-мишеней нами были использованы сведения, представленные в обучающей выборке PASS Targets. На данном этапе, мы поставили цель сформировать набор белков-мишеней так, чтобы он наилучшим образом мог охарактеризовать химическое соединение и, поэтому в его состав вошли мишени, для которых представлены наилучшие значения инвариантной точности прогноза.

Т.И. Маджидов: Проверялось ли, что при предсказании активности молекулы, она не входила в обучающую выборку PASS Targets? Учитывая большой объем обучающей выборки, использованной PASS это вполне возможно. Если они все были в обучении, то есть шанс, что предсказания будут повторять экспериментальный профиль активности. И в этом случае это будет эквивалентно кластеризации по экспериментальным активностям. Кстати, не пробовали ли это?

Н.С. Ионов: В исследованном наборе соединений, соединения которые могут быть обнаружены в обучающей выборке PASS Targets, для сформированного нами набора белков-мишеней, представлено не было. Кластеризация по экспериментальным активностям не проводилась.

Т.И. Маджидов: Имелся ли какой-то смысл у двух выделенных кластеров? Скажем в одном группировались соединения с множественной активностью, а в другом - с меньшим промискуитетом? Или может группировка была по группа биоактивности?

Н.С. Ионов: Анализ такого рода проводился, однако существенных особенностей, при разделении синтетических и природных соединений на два кластера, нами выявлено не было.