

А.А. Рудик: Какой был размер выборки (корпуса Chemnder), использованный при построении модели?

Н.Ю. Бизюкова: Для построения модели использовалась часть корпуса CHEMDNER (обучающая выборка) в размере 1000 абстрактов.

После публикации презентации доклада мы расширили эту выборку до 2500 и провели дополнительную валидацию, и при этом точность прогноза существенно увеличилась.

	Precision	Recall	F1-score
S	0,88	0,72	0,79
O	0,98	0,99	0,99
B	0,89	0,82	0,86
I	0,96	0,97	0,96
E	0,88	0,81	0,84
Среднее	0,92	0,82	0,89

Precision: 92,02

Recall: 82,27

F1-score: 88,92