

Чекпоинт 5
нелинейные ML-модели
классификатор пресс-релизов ЦБ РФ
команда 14

Произведено обучение и перебор по сетке гиперпараметров для следующих нелинейных моделей:

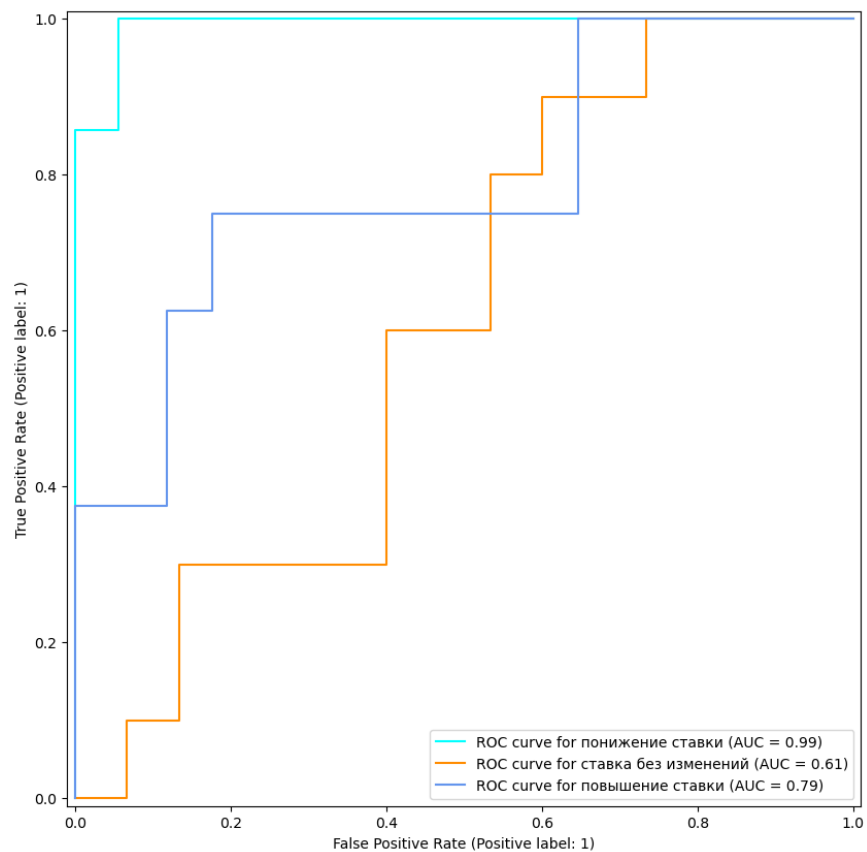
- 1) решающее дерево;
- 2) случайный лес;
- 3) CatBoost.

CatBoost смог показать качество чуть выше лучшей линейной модели.

Сравнение нелинейных моделей с лучшей линейной

	f1-score «-1»	f1-score «0»	f1-score «1»	accuracy	auc-score
DecicionTreeClassifier N-GRAM (1, 3) TF-IDF	0.8	0.35	0.0	0.4	-
RandomForestClassifier N-GRAM (1, 3) TF-IDF	0.71	0.48	0.53	0.56	-
CatBoostClassifier N-GRAM (1, 3) TF-IDF	0.67	0.62	0.5	0.6	0.86
LogReg на TF-IDF N-GRAM (1, 2, 3)	0.36	0.67	0.67	0.6	0.86

Log Geg N-GRAM tfidf ROC-AUC



CatBoost N-GRAM tfidf ROC-AUC

