

### 0.1. *Мадьяров К.* МЛ-конвейер для прогнозирования и выявления погодных причин аварийных ситуаций на линиях электропередач

Выявление причин и прогнозирование сбоев объектов критической инфраструктуры, в том числе линий электропередач, является важной задачей не только для обеспечения их стабильной работы, но и для повышения эффективности деятельности страховых компаний.

При использовании методов машинного обучения для решения данной задачи в настоящее время можно выделить две особенности. Во-первых, при оценках эффективности прогнозирования зачастую специалисты ограничиваются только метриками качества классификации, такими как accuracy, precision, recall, F1-мера и ROC-AUC. Однако для сопоставления моделей в условиях реальной эксплуатации требуется важно учитывать и ресурсные показатели, включающие время обучения и пиковое потребление ОЗУ. Во-вторых, могут встречаться «засоры» и неоднородности датасетов, используемых для обучения, что может приводить к снижению качества моделей и появлению ложных связей.

Для решения обозначенных проблем в настоящей работе предлагается схема построения МЛ-конвейера, включающего модель машинного обучения с автоматизированным подбором гиперпараметров в Optuna и интерпретацию результатов посредством SHAP-анализа [1], позволяющую выявить ключевые погодные факторы риска (скорость ветра, атмосферное давление и относительную влажность). Для устранения дисбаланса классов применено взвешивание классов (модель уделяет больше внимания редким событиям) вместо синтетического увеличения выборки.

МЛ-конвейер апробирован на открытом датасете из Kaggle, в котором применены данные об энергокомплексе Краснодарского края [2]. Для обучения конвейера объединены открытые архивные данные об авариях ЛЭП (2012–2019 гг.) и почасовые метеоданные восьми станций. Числовые и категориальные (выраженные не числом, а меткой) агрегированы до временного шага 1 час. Данные характеризуются сильным дисбалансом классов, поскольку аварийные ситуации достаточно редки.

Эксперименты показали, что наибольшую эффективность МЛ-конвейер дает при использовании модели CatBoos, достигшей наивысшего ROC-AUC приблизительно 0,71 при полноте приблизительно 0,61 на тестовых данных (базовый порог 0,5). При этом CatBoost продемонстрировал и наилучшее соотношение качество/ресурсы.

*Научный руководитель — к.ф.-м.н. Пестунов А. И.*

#### Список литературы

[1] SCOTT LUNDBERG SHAP latest documentation. URL: <https://shap.readthedocs.io/en/latest/> (дата обращения:

11.05.2024).

[2] СУСЛОВ А. Accidents on the energetical objects [Электронный ресурс] // Kaggle. – URL: <https://www.kaggle.com/code/suslovaalexey/accidents-on-the-energetical-objects> (дата обращения: 10.09.2024).