

Разработка программного продукта для анализа и прогнозирования концентрации газов в трансформаторном масле силовых трансформаторов

А.Ю.Белякова¹, Ю.Д.Беляков²

*1Иркутский государственный аграрный университет имени А.А. Ежевского,
Иркутск*

2Санкт-Петербургский государственный университет, Санкт-Петербург

Аннотация: Статья посвящена разработке программного продукта, анализирующего концентрацию газов (H₂; CO; C₂H₄; C₂H₂), измеренных в трансформаторном масле и заданных в виде зависимости от времени. Разработка продукта состоит из двух этапов. Первый этап заключается в задаче классификации трансформаторов. Второй этап - прогнозирование надежной работы до возможного отказа по заданным в качестве исходных данных концентрациям растворенных газов в трансформаторном масле. При разработке программного продукта были использованы фреймворки для создания веб-приложений Flask и ReactJS, модели обучены на фреймворках H2O.ai, LightGBM, Sklearn.
Ключевые слова: машинное обучение, регрессия, силовые трансформаторы, концентрация газов, веб-сервис.

Постановка задачи

На сегодняшний день значительное количество силовых трансформаторов АЭС (блочные трансформаторы, трансформаторы собственных нужд, автотрансформаторы связи) (далее – силовые трансформаторы) эксплуатируются с продленным сроком службы (назначенный срок службы – 25 лет). Принимая во внимание продленные сроки эксплуатации энергоблоков АЭС, возникает потребность в контроле за техническим состоянием силовых трансформаторов.

Особую роль в контроле за маслосодержащем оборудованием играет контроль газов, растворенных в масле [1]. Появление в оборудовании дефектов сопровождается образованием газов, при этом специфические виды дефектов генерируют свои газы в разных количествах. Однако, существующие способы идентификации дефекта, основанные на измерении газов, работают только при условии достижения концентрациями газов некоторой установки [2].

Задача данной работы - по временным рядам предсказать время, через которое концентрации газов достигнут установки (либо максимальное время прогноза, если установка не будет достигнута).

Решение поставленной задачи

В работе рассмотрены две задачи: первая - классифицировать трансформаторы на исправные и неисправные, а затем определить вид неисправности. Таким образом, трансформатор может относиться к одному из 4-х классов: исправная работа, частичный разряд, разряд низкой энергии и низкотемпературный термический дефект. И вторая – предсказывать, через какое время концентрации газов достигнут установки. Каждому трансформатору в соответствие было поставлено число от 362 до 1093, обозначающее количество временных промежутков, через которые будут достигнуты установки.

Решения данных задач сводятся к формированию признаков, которые наилучшим образом описывают временные ряды, и к выбору наилучшей модели машинного обучения [3, 4]. Для этого была представлена обучающая выборка, состоящая из описаний состояний более двух тысяч силовых трансформаторов.

Обучающие данные представлены в виде временных рядов. Выборка состоит из 2100 файлов, в каждом из которых на промежутке в 420 единиц времени отображены концентрации 4-х газов (H₂; CO; C₂H₄; C₂H₂).

Таблица № 1

Формат файлов с данными о концентрации газов

	H ₂	CO	C ₂ H ₄	C ₂ H ₂
0	0,001107	0,026212	0,006355	0,000021
1	0,001107	0,026210	0,006353	0,000021
2	0,001107	0,026208	0,006350	0,000021
3	0,001107	0,026207	0,006348	0,000021
4	0,001107	0,026206	0,006346	0,000021

В задаче классификации метка для каждого файла - одни из 4-х классов. В задаче регрессии - количество временных промежутков.

На первом этапе решения данной проблемы были выбраны признаки. Чем лучше признаки - тем лучше модель.

Опираясь на тренировочные данные и изучив уже существующие методические материалы (СТО 34.01-23-003-2019), сформирован список признаков, который включает в себя:

- во-первых, статистические показатели временных рядов (это средние, максимальные, минимальные значения концентраций, стандартное отклонение и корреляции между рядами);
- во-вторых, отношения концентраций (это отношения средних, максимальных, минимальных и последних значений);
- в-третьих, показатели, связанные со скоростью изменения концентраций (это скорости изменения и количества достигаемых экстремумов);
- в-четвертых, концентрации газов в процентном соотношении;
- и в-пятых, для передачи информации о трендах роста концентраций, мы использовали отношения концентраций газов к их допустимым значениям в нескольких точках.

Все эти признаки использовались как для тренировки модели классификации, так и модели регрессии.

Второй этап - это выбор моделей. Был проведен большой анализ различных методов машинного обучения в самых разных конфигурациях и комбинациях. В результате, наилучшие показатели, как для задачи классификации, так и для регрессии, показывали модели, основанные на решающих деревьях, а именно: LightGBM [5], Extra Trees [6], Random Forest [7] и Extreme Gradient Boosting [8] (табл.2). При анализе были

использованы следующие метрики машинного обучения: Accuracy, AUC, Recall, Precision, f1-score.

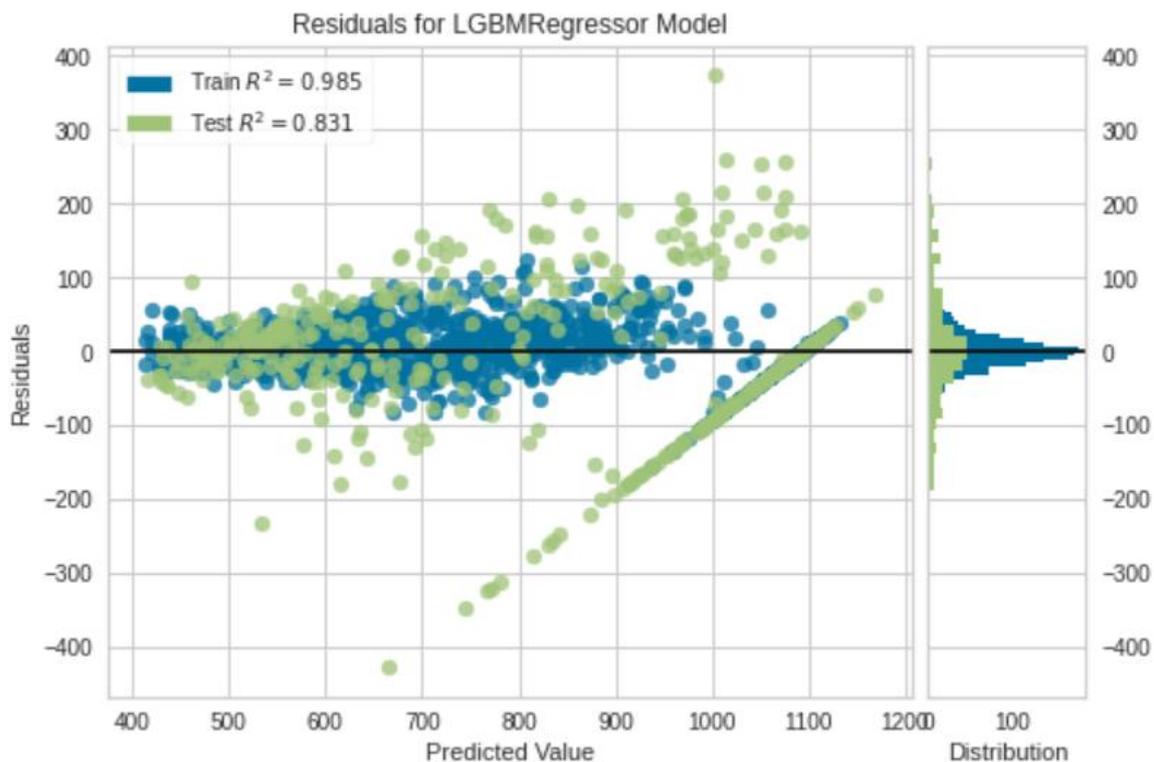
Таблица № 2

Показатели моделей в задаче классификации

Model	Accuracy	AUC	Recall	Prec.	F1
Light Gradient Boosting Machine	0.9768	0.9973	0.9319	0.9775	0.9763
Extreme Gradient Boosting	0.9726	0.9977	0.9215	0.9736	0.9723
Random Forest Regressor	0.9714	0.9978	0.9206	0.9721	0.9711
Gradient Boosting Regressor	0.9714	0.9967	0.9230	0.9722	0.9711
Extra Trees Regressor	0.9714	0.9975	0.9143	0.9722	0.9706
Decision Tree Classifier	0.9542	0.9402	0.8682	0.9549	0.9534
Linear Discriminant Analysis	0.9482	0.9787	0.8342	0.9497	0.9446
Ridge Classifier	0.9238	0.0000	0.7675	0.9263	0.9104
K Neighbors Classifier	0.8958	0.9553	0.6646	0.8852	0.8866
Quadratic Discriminant Analysis	0.8946	0.9375	0.6052	0.8942	0.8833
Ada Boost Classifier	0.8875	0.9069	0.7310	0.9119	0.8913
Logistic Regression	0.8542	0.9074	0.4878	0.7867	0.8084
SVM Linear Kernel	0.8321	0.0000	0.4099	0.7593	0.7872
Naïve Bayes	0.0595	0.6816	0.2543	0.7104	0.0306

В задаче регрессии все обстояло не так просто. После того, как модель обучили на всех данных, обнаружили, что на графике невязок есть четкий паттерн - линия. Это связано с тем, что у меток временных рядов есть верхняя граница - 1093. Эта метка ставится тогда, когда установка будет достигнута через очень долгое время, а значит метки 1093 и, например, 1092 различаются всего на 1 временной промежуток, но в реальности это могут быть сотни или тысячи промежутков.

Это значительно сказывается на интерпретации признаков (рис.1).



Рис

. 1. – График невязок

При этом, значение 1093 - самое часто встречающееся (33% всех измерений) (рис.2).

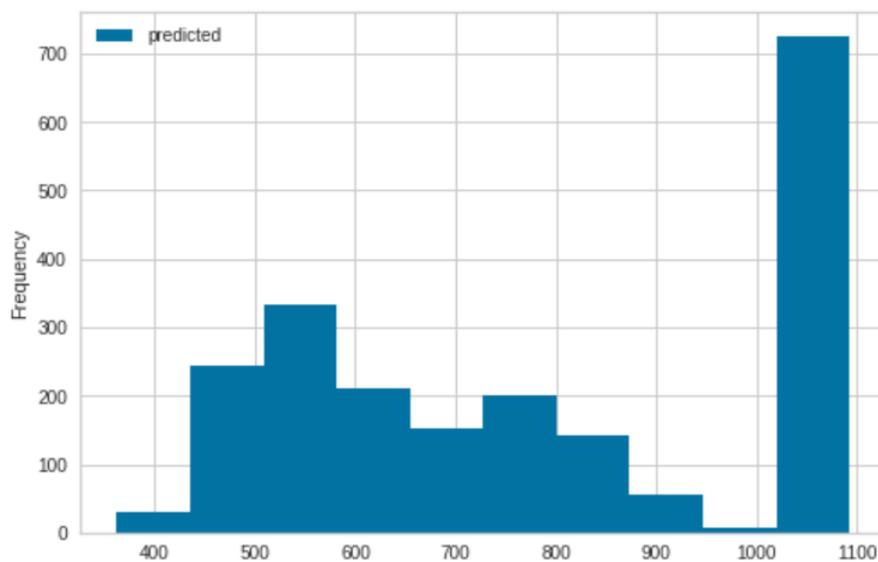


Рис. 2. - Распределение элементов обучающей выборки по временным меткам

Чтобы избавиться от этого, задачу разбили на две части. Сначала выборка разбивается на два подмножества: элементы с временной меткой 1093 и элементы с временной меткой <1093. Далее решается задача бинарной классификации: показатели достигнут установки до 1093 временных промежутков или после. И затем задача регрессии: предсказание количества временных промежутков для тех показателей, которые были классифицированы как <1093. Показатели моделей при обучении с новым подходом отображены в (табл.3).

Таблица №3

Показатели моделей в задаче регрессии

Model	MAE	MSE	RMSE	R2
Light Gradient Boosting Machine	49,1389	4868.5587	69.5309	0.7621
Extra Trees Regressor	53,7586	5364.3173	73.0136	0.7407
Random Forest Regressor	52,1948	5478.8383	73.7573	0.7331
Extreme Gradient Boosting	55.3061	5709.7646	75.4133	0.7225
Gradient Boosting Regressor	59.6302	6241.8021	78.8851	0.6959
Ada Boost Classifier	84.3864	10104.5394	100.2531	0.5100
Decision Tree Regressor	71.2951	11068.0819	104.6438	0.4585
Orthogonal Matching Pursuit	99.1040	15078.1734	122.5324	0.2707
Lasso Least Angle Regression	122.2381	20571.5751	143.2080	0.0054
Bayesian Ridge	124.8175	23280.2936	150.9391	-0.1118
K Neighbors Regressor	126.1842	23371.5963	152.6417	-0.1317

На таком обучающем множестве “без 1093” получены результаты значительно лучшие, относительно предыдущей модели и на графике невязок уже не заметно таких четких паттернов (рис 3.).

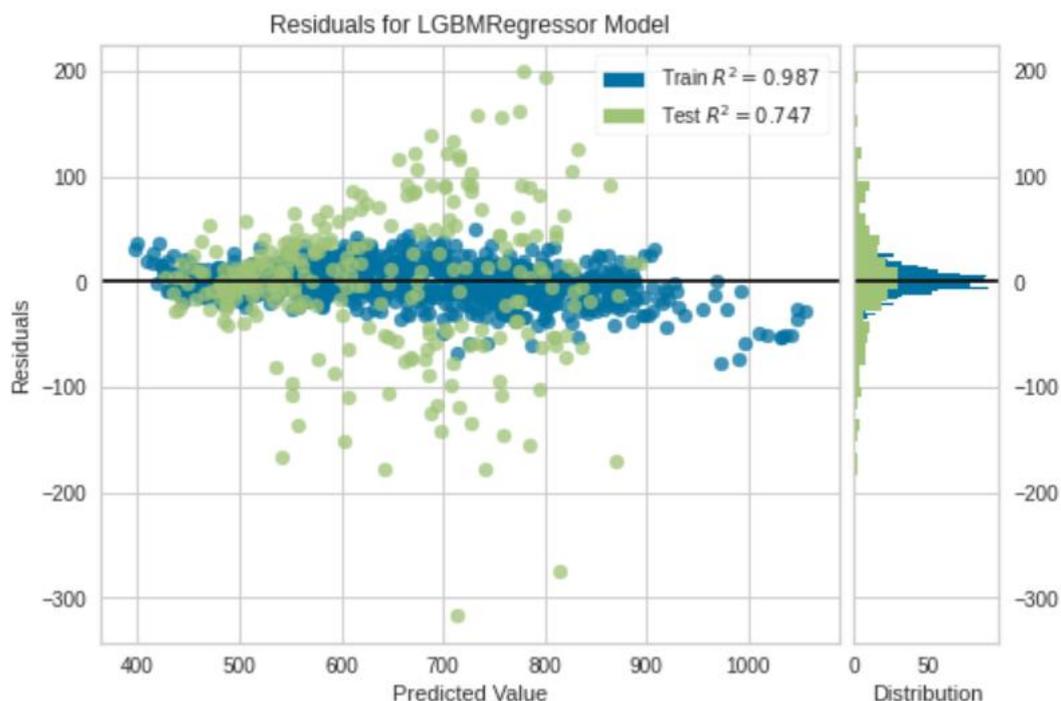


Рис. 3. – График невязок после изменения модели

Результаты

Разработанные модели предлагается использовать для создания веб-сервиса [9,10]. С его помощью пользователь может контролировать техническое состояние силовых трансформаторов, следить за текущими концентрациями газов в масле, узнавать о возможных поломках, а также получать прогноз о достижении установок.

На рисунке 4 представлен стек использованных технологий при разработке:

- фреймворки веб- сервисов Flask и ReactJS;
- модели были обучены на фреймворках H2O.ai, LightGBM, Sklearn.



Рис. 4. – Стек использованных технологий

Веб-сервис имеет следующую архитектуру (рис.5). База данных обновляется информацией с онлайн газоанализаторов. С помощью модели машинного обучения новые данные обрабатываются и формируются обновленные предсказания. На веб-сервисе пользователь контролирует технические состояния трансформаторов, используя полученные предсказания.

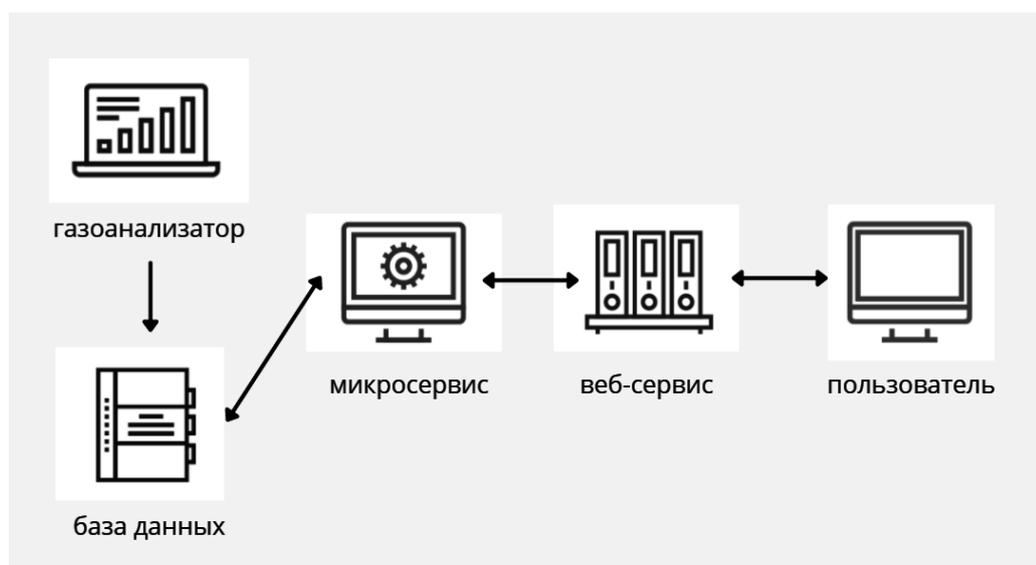


Рис. 5. - Архитектура веб-сервиса

На рисунке 6 показана демонстрация работы приложения. В выпадающем меню можно выбрать необходимый трансформатор и узнать его состояние на данный момент.



Рис. 6. – Демонстрация окна приложения

Например, в первом трансформаторе наблюдается разряд низкой энергии, а концентрации его газов достигнут установки через 721 день. Также можно посмотреть показатели различных газов в масле.

Выводы

Разработан программный продукт для анализа текущего состояния силовых трансформаторов и прогнозирования надежной работы до возможного отказа по измеренным за время мониторинга концентрациям растворенных газов в трансформаторном масле газоанализаторами.

Разработанный продукт можно улучшить увеличением объема данных. Также необходимо провести еще более глубокий анализ моделей и создать среду для дообучения моделей.

Литература

1. Овсянников Ю.М., Попов Г.В. Прогнозирование состояния силового трансформатора по результатам хроматографического анализа растворённых газов // Известия высших учебных заведений. Электромеханика. 2015. №1. С. 53-60. URL: elibr.ary.ru/item.asp?id=23025020
2. Туркин Д.Г., Волокитин Р.А. Новые тенденции в контроле технического состояния электроэнергетического оборудования // Международный научно-исследовательский журнал. 2020. №10 С. 81—84. URL: research-journal.org/technical/novye-tendencii-v-kontrole-texnicheskogo-sostoyaniya-elektroenergeticheskogo-oborudovaniya
3. Болдырев С.В. Применение гибридных самоорганизующихся нейронных сетей и быстрого дискретного вейвлет-преобразования для построения систем классификации сигналов. // Инженерный вестник Дона. 2012. №2. URL: ivdon.ru/ru/magazine/archive/n2y2012/773
4. Алёшин С.П., Бородина Е.А. Нейросетевое распознавание классов в режиме реального времени // Инженерный вестник Дона. 2013. №1. URL: ivdon.ru/ru/magazine/archive/n1y2013/1494
5. Ke, Guolin et al. “LightGBM: A Highly Efficient Gradient Boosting Decision Tree.” // Conference on Neural Information Processing Systems (NIPS). 2017. URL: papers.nips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
6. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. // Mach Learn. 2006. 33–42. URL: doi.org/10.1007/s10994-006-6226-1
7. Breiman, L. Random Forests. Machine Learning. 2001. 5–32. URL: doi.org/10.1023/A:1010933404324
8. Chen, T., & Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings // 22nd ACM SIGKDD International Conference on Knowledge

Discovery and Data Mining New York, NY, USA ACM. 2016. 785–794. URL: doi.org/10.1145/2939672.2939785

9. Лиля В.Б., Костюков А.В. Экспертная система диагностики силовых трансформаторов // Инженерный вестник Дона. 2013. №1. URL: ivdon.ru/ru/magazine/archive/n1y2013/1504

10. Натальченко И.А. Анализ механизмов передачи крупных массивов данных через сеть интернет с помощью технологии веб-сервиса // Инженерный вестник Дона. 2008. №4. URL: ivdon.ru/ru/magazine/archive/n4y2008/98

References

1. Ovsjannikov Ju.M., Popov G.V. Izvestija vysshih uchebnyh zavedenij. Jelektromehanika. 2015. №1. S. 53-60

2. Turkin D.G. Volokitin, R.A. Mezhdunarodnyj nauchno-issledovatel'skij zhurnal. 2020. №10. pp. 81—84.

3. Boldyrev S.V. Inzhenernyj vestnik Dona. 2012. №2. URL: ivdon.ru/ru/magazine/archive/n2y2012/773

4. Aljoshin S.P., Borodina E.A. Inzhenernyj vestnik Dona. 2013. №1. URL: ivdon.ru/ru/magazine/archive/n1y2013/1494

5. Ke, Guolin et al. Conference on Neural Information Processing Systems (NIPS).2017.URL:papers.nips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

6. Geurts, P., Ernst, D. & Wehenkel, L. Mach Learn. 2006. 33–42. URL: doi.org/10.1007/s10994-006-6226-1

7. Breiman, L. 2001. 5–32. URL: doi.org/10.1023/A:1010933404324

8. Chen, T., & Guestrin, C. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining New York, NY, USA ACM. 2016. 785–794. URL: doi.org/10.1145/2939672.2939785

9. Lila V.B., Kostjukov A.V. Inzhenernyj vestnik Dona. 2013. №1. URL:
ivdon.ru/ru/magazine/archive/n1y2013/1504

10. Natal'chenko I.A. Inzhenernyj vestnik Dona. 2008. №4. URL:
ivdon.ru/ru/magazine/archive/n4y2008/98