

ДИАГНОСТИРОВАНИЕ СОСТОЯНИЯ ШТАНГОВЫХ ГЛУБИННЫХ НАСОСОВ ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ

Аннотация. В статье представлено описание процесса разработки модели нейронной сети для решения задачи распознавания состояния работы штангового глубинного насоса (ШГН) по динамограмме. Проведены сравнения полученного решения с ранее опубликованными. Сделаны выводы о направлении дальнейших исследований в данной области.

Ключевые слова: разметка данных, машинное обучение, штанговый глубинный насос, скважина, динамограмма, контроль, динамометрирование, нейронные сети.

Введение. Нефтегазовая отрасль является одной из ключевых отраслей мировой экономики, где сбор, мониторинг, своевременное реагирование на состояние оборудования играют важнейшую роль для принятия решений по повышению производительности скважин и предотвращению возможных аварийных ситуаций, от чего зависит эффективность и стабильность работы скважин и месторождений в целом.

Установки штанговых глубинных насосов (ШГН) остаются наиболее распространенным методом добычи нефти. Типовым приводом ШГН является станок-качалка, который преобразует вращательное движение вала двигателя в возвратно-поступательное движение точки повеса штанг. В настоящее время такими установками оснащены более половины скважин на территории России, а в некоторых странах их количество достигает до 90%. Наиболее эффективным из способов диагностирования состояния ШГН остается динамометрирование, — построение динамограммы. Динамограмма отражает зависимость нагрузки в точке подвеса штанг от положения полированного штока, которая дает информацию о состоянии глубинного оборудования скважины [1].

Диагностирование состояния глубинного оборудования обычно проводится группой экспертов — 2-3 человека. В один рабочий день продолжительностью 8 часов, эксперт может составить заключение о состоянии оборудования примерно по 1000 динамограммам. Пример динамограммы представлен на рис. 1. По оси X — положения полированного штока, Y — нагрузка в точке подвеса штанг.

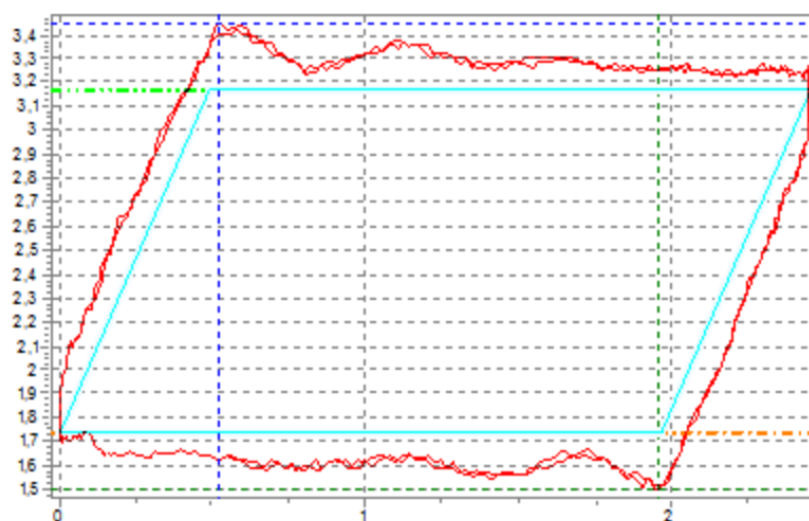


Рис. 1. Пример динамограммы

В процессе диагностирования состояния ШГН можно выделить следующие проблемы:

- большие человеческие и временные затраты на диагностирование состояния погружного оборудования одной скважины;
- наличие человеческого фактора. Эксперты могут допускать ошибки, которые приводят к расхождениям оценок экспертов между собой, что отсрочит принятие необходимого решения по эксплуатационному режиму оборудования, а как следствие — снижение производительности добычи;
- сложность и монотонность процесса. Процесс анализа динамограмм требует концентрации и внимательности от эксперта, чтобы верно составить заключение по состоянию оборудования на основе динамограммы.

Безусловно, данные проблемы уже решались другими исследователями, но как правило в работах используется ограниченный список диагностируемых состояний, который регламентируется нормативной документацией компаний.

Из некоторых работ тезисно были выделены следующие примеры перечня состояний:

1. Определяется 3 типовых класса состояния ШГН, отсутствует детекция одновременных проявлений нескольких состояний насоса [2].
2. Используется довольно широкий перечень классов состояний ШГН, помимо типовых, но не все состояния актуальны для отечественных насосов, а также отсутствует детекция нескольких состояний одновременно, но некоторые представляют собой комбинации нескольких состояний [3].
3. Используется перечень из 10 типовых классов состояний насоса, также отсутствует детекция нескольких состояний одновременно [4].
4. Используется 7 типовых классов состояния насосов с использованием CV [5].

Проблема исследования. Основной проблемой в процессе диагностировании состояния ШГН — высокая вовлеченность людей в рутинную, монотонную и требующей концентрации работу [6]. В связи с описанными ранее недостатками диагностики состояния ШГН, возникает потребность в создании инструментов, позволяющих автоматизировать процесс диагностирования.

Материалы и методы. Исходя из описания процесса диагностики состояния ШГН, в первую очередь необходимо определить перечень состояний, которые будут классифицироваться моделью, а также способ разметки состояний. В процессе изучения работ других исследователей и нормативной документации было замечено, что в большинстве случаев регламентируемый перечень состояний является комбинациями более узких факторов. На основе этого был определен перечень факторов состояний, по которым диагностируются состояния ШГН.

Используя различные комбинации факторов, можно определить большинство возможных регламентированных состояний ШГН. Иными словами, вместо конкретных классов состояний модель использует факторы, которые указывают на их наличие проявления состояния. Это позволяет модели распознавать одновременное проявление нескольких состояний работы насоса. Отсутствие факторов интерпретируется как нормальная работа ШГН.

Первоочередной задачей и самой главной для разработки для модели — сбор обучающего набора. При удалось получить датасет объемом в 30 000 записей динамограмм, который

представляет собой последовательность точек X и Y различной длины, который был обработан стандартными методами интерполяции для приведения всех наборов одинаковой размерности.

Немаловажным этапом стала разметка датасета. Факторы в обучающей выборке были представлены в формате one-hot encoding. По данному формату было размечено 6000 динамограмм.

При исследовании получившегося датасета были получены следующие результаты:

Как и предполагалось, некоторые факторы имеют довольно высокий показатель корреляции;

Представителей некоторых факторов довольно мало (табл. 1) в получившемся датасете, что может стать проблемой при обучении модели нейронной сети.

Разработка архитектуры и модели. В данном разделе мы рассмотрим основные этапы разработки нейронной сети и варианты конфигураций для удовлетворительных результатов, которые мы предприняли.

Определение архитектуры

Первым шагом является определение архитектуры нейронной сети. Это включает в себя выбор типа сети количество скрытых слоев и количество нейронов в каждом слое. В зависимости от задачи, требуемой точности и ресурсов можно рассмотреть различные варианты архитектуры. На первой итерации мы остановились на полносвязной нейронной сети, с входным слоем в 1000 нейронов и тремя скрытыми слоями

Выбор функций активации. Эксперименты с различными функциями активации могут помочь определить подходящий вариант для конкретной задачи. Для выходного слоя, в разработанной модели используется Sigmoid-ая функция активации, так как перед нами стоит задача мультиклассовой классификации, для всех остальных слоев было принято решение использовать функцию активации ReLU.

Оптимизация параметров. После определения архитектуры и функций активации необходимо подобрать оптимальные параметры обучения, такие как скорость обучения и метод оптимизации. Это может включать в себя использование методов градиентного спуска, адаптивных оптимизаторов и регуляризацию для предотвращения переобучения. В нашем случае было принято решение остановиться на 150-ти эпохах обучения и использовать в качестве функции потерь Бинарную Кроссэнтропию, в качестве оптимизатора: Adam, и метрика качества Accuracy.

Результаты. После определения рабочей конфигурации модели производится ее финальная настройка, включая регуляризацию, настройку гиперпараметров и улучшение обобщающей способности модели. Для нас такими настройками стали: проверка данных, тестирование обучения на разном количестве эпох, и тестирование других различных методов.

В ходе обучения выбранной конфигурации модели, на валидационной выборке были получены следующие показатели качества, которые представлены на рис. 2.

Анализируя показатели, видим, что точность модели довольно низкая, но это обусловлено тем, что при подходе, который использовался в работе, нас интересует классификация наличия факторов, а не однозначная классификация состояния. Результаты можно интерпретировать следующим образом — для 50% динамограмм полный перечень факторов классифицируется верно, что является хорошим результатом, учитывая несбалансированность выборки по различным факторам.

	precision	recall	f1-score	support
0	0.77	0.70	0.73	296
1	0.46	0.21	0.29	87
2	0.38	0.24	0.29	58
3	0.24	0.22	0.23	50
4	0.26	0.26	0.26	50
5	0.40	0.45	0.42	100
6	0.55	0.66	0.60	341
7	0.26	0.45	0.33	87
8	0.22	0.19	0.20	72
9	0.00	0.00	0.00	5
10	0.33	0.12	0.18	8
11	1.00	0.20	0.33	5
12	0.00	0.00	0.00	15
14	0.67	0.57	0.62	7
15	0.00	0.00	0.00	1
16	0.50	0.25	0.33	4
17	1.00	0.20	0.33	5
18	0.00	0.00	0.00	3
19	1.00	0.33	0.50	6
accuracy			0.50	1200
macro avg	0.42	0.27	0.30	1200
weighted avg	0.50	0.50	0.49	1200

Рис 2. Показатели точности модели на тестовой выборке

Заключение. В ходе разработки модели распознавания состояния ШГН была получен перспективный подход к решению поставленной задачи, но, к сожалению, оно еще требует доработок. Одними из главных достоинства модели можно считать следующие аспекты:

- минимальные преобразования исходных данных для взаимодействия с моделью;
- скорость обучения модели;
- скорость работы модели.

Основным недостатком текущей модели является ее недостаточная точность, составляющая всего 50% по метрике Ассигасу. Нефтяные компании, занимающиеся добычей, заинтересованы в высокой точности моделей. Использование модели с точностью в 50% сопряжено с существенными рисками для их предприятий.

В сравнении с работами других исследователей, сообщающих о точности в диапазоне 70-90%, следует отметить, что в этих исследованиях перечень состояний строго регламентирован, и нет возможности классифицировать одновременное проявление нескольких состояний. Вероятно, именно из-за этого различия в подходах модель в текущем исследовании имеет более низкую точность. И важно отметить, что в работах для обучения и тестирования модели используются теоретические динамограммы (расчетные), а в данной работе обучение и тестирование проводится на реальных данных, которые помимо наличия шумов и выбросов, не сбалансированы по количеству представителей факторов.

Есть несколько вариантов устранения данного недостатка, а также вариантов апробации модели в различных кейсах:

1. Расширение обучающего набора данных, в ходе которого будет устранена проблема несбалансированности данных в обучающем наборе.
2. Undersampling — уменьшения количества примеров динамограмм с более представленными факторами, для сбалансирования датасета.
3. Использование CV (Computer Vision).

СПИСОК ЛИТЕРАТУРЫ

1. Бубнов М.В. Средства диагностирования оборудования установок штанговых глубинных насосов / М.В. Бубнов, А.М. Зюзев // Труды первой научно-технической конференции молодых ученых Уральского энергетического института, Екатеринбург, Россия, 16-20 мая 2016 г. — Екатеринбург: [УрФУ], 2016. — С. 175-178.
2. Yahar Ghasemi Habashi Application of Neural Networks to Diagnostics of Pumping-Rod Units // IV International Conference “Problems of Cybernetics and Informatics” (PCI’2012), September 12-14, 2012. — Baku: IEEE, 2012. — С. 259-260. 12.09.2012. (дата обращения: 04.05.2024)
3. Patrick Bangert Diagnosing and Predicting Problems with Rod Pumps Using Machine Learning // In Proceedings of the SPE Middle East Oil and Gas Show and Conference. — Manama: SPE, 2019 (дата обращения: 04.05.2024).
4. Yun-Peng He, Chuan-Zhi Zang, Peng Zeng, Ming-Xin Wang, Qing-Wei Dong, Guang-Xi Wan, Xiao-Ting Dong Few-shot working condition recognition of a sucker-rod pumping system based on a 4-dimensional time-frequency signature and meta-learning convolutional shrinkage neural network // Petroleum Science. — 2023. — №20. — С. 1142-1154 (дата обращения: 04.05.2024).
5. Yun-Peng He, Chuan-Zhi Zang, Peng Zeng, Ming-Xin Wang, Qing-Wei Dong, Guang-Xi Wan, Xiao-Ting Dong Working condition recognition of sucker rod pumping system based on 4-segment time-frequency signature matrix and deep learning // Petroleum Science. — 2024. — № 21. — С. 641-653 (дата обращения: 04.05.2024).
6. Хакимьянов М.И. Опыт диагностирования состояния глубиннонасосного оборудования системой динамометрирования ДДС-04: автореф. дис. ... канд. техн. наук / М.И. Хакимьянов, С.В. Светлакова, М.Е. Сидоров. — Уфа, 2008. — С. 7.