

Методология автоматизированной подготовки данных для машинного обучения нейросетевых моделей в интеллектуальных системах выявления и прогнозирования осложнений и аварийных ситуаций в процессе строительства нефтяных и газовых скважин

Еремин Н.А., Черников А.Д.

Институт проблем нефти и газа РАН, Москва, Россия
ermn@mail.ru, cha60@mail.ru

Аннотация

В статье описана методология автоматизированной обработки промышленной информации и подготовки данных для обучения нейросетевых моделей в интеллектуальных системах выявления и прогнозирования прихватов буровых колонн, применяемых для повышения безопасности и сокращения непроизводительного времени при строительстве нефтяных и газовых скважин. Предлагаемая методика позволяет проводить автоматизированный анализ больших объемов архивной информации (Big Geo Data), выделять характерные ситуации, связанные с осложнениями, и осуществлять разметку информационных массивов данных для машинного обучения унифицированных интеллектуальных модулей обработки данных от станций геолого-технологических измерений с целью заблаговременного предупреждения операторов бурового оборудования о технологических рисках возникновения нестандартных ситуаций. Описаны методологические подходы к разметке данных для одноклассовых и многоклассовых нейросетевых моделей прогнозирования возникновения осложнений при строительстве нефтяных и газовых скважин на суше и на море.

Материалы и методы

В статье использованы результаты исследований и опытной эксплуатации экспериментального образца автоматизированной системы предупреждения осложнений и аварийных ситуаций при строительстве скважин (АС ПОАС), разработанного в ИПНГ РАН в рамках поисковой научной работы по заказу Министерства образования Российской Федерации. Применялись методы автоматизации подготовки и детальной разметки данных для построения и применения многоклассовых нейросетевых моделей прогнозирования возникновения осложнений при строительстве нефтяных и газовых скважин.

Ключевые слова

одноклассовая нейросетевая модель прогнозирования, многоклассовая нейросетевая модель прогнозирования, осложнения и аварийные ситуации при строительстве скважин, прихваты буровой колонны, данные геолого-технологических измерений, автоматизированная обработка промышленной информации, машинное обучение, нейросетевые модели, универсальные интеллектуальные модули и комплексы

Статья подготовлена в рамках выполнения государственного задания, номер гос. рег. № НИОКТР в РОСРИД 122022800270-0.

Для цитирования

Еремин Н.А., Черников А.Д. Методология автоматизированной подготовки данных для машинного обучения нейросетевых моделей в интеллектуальных системах выявления и прогнозирования осложнений и аварийных ситуаций в процессе строительства нефтяных и газовых скважин // Экспозиция Нефть Газ. 2024. № 5. С. 24–30. DOI: 10.24412/2076-6785-2024-5-24-30

Поступила в редакцию: 16.08.2024

DRILLING

UDC 622.276.66 | Original Paper

Methodology of automated data preparation for machine learning of neural network models in intelligent systems for identifying and predicting complications and emergency situations during the construction of oil and gas wells

Eremin N.A., Chernikov A.D.

Institute of Oil and Gas Problems of the Russian Academy of Sciences, Moscow, Russia
ermn@mail.ru, cha60@mail.ru

Abstract

Research Topic: Methodology for Automated Data Preparation for Machine Learning of Neural Network Models in Intelligent Systems for Detecting and Predicting Complications and Emergencies in the Construction of Oil and Gas Wells.

The article presents a methodology for automated processing of field information and data preparation for training neural network models in intelligent systems for detecting and predicting drill string sticking, used to improve safety and reduce non-productive time in the construction of oil and gas wells. The developed methodology allows for automated analysis of large volumes of archival information, identification of typical situations associated with complications, and marking of information arrays for machine learning of unified intelligent modules for processing data from geological and technological measurement stations in order to provide early warning to drilling equipment operators about technological risks of emergency situations.

Materials and methods

The article uses the results of research and trial operation of an experimental prototype of an automated system for preventing complications and emergency situations during well construction, developed at the Institute of Oil and Gas Problems of the Russian Academy of Sciences as part of exploratory research commissioned by the Ministry of Education of the Russian Federation. Methods of automating the preparation and detailed labeling of data were used to construct and apply multi-class neural network models for predicting

For citation

Eremin N.A., Chernikov A.D. Methodology of automated data preparation for machine learning of neural network models in intelligent systems for identifying and predicting complications and emergency situations during the construction of oil and gas wells. Exposition Oil Gas, 2024, issue 5, P. 24–30. (In Russ). DOI: 10.24412/2076-6785-2024-5-24-30

Received: 16.08.2024

Российская нефтегазовая экономика ищет новые точки роста на пути цифровой модернизации производства и экономики больших высокочастотных геоданных. Нефтегазовые компании с эффективной моделью управления, высоким уровнем цифровизации и высококвалифицированным персоналом могут обеспечить рост объемов собственного производства примерно в 2–3 раза [1]. Технологии искусственного интеллекта используются при оптимизации производственных процессов и автоматизации бурения. Особое значение имеет использование интеллектуальных систем для прогнозирования и предотвращения осложнений и аварийных ситуаций при строительстве скважин [1]. Основным методом машинного обучения систем ИИ, применяемых в нефтегазовой отрасли, является «обучение с учителем» [2]. В качестве «учителей» выступают эксперты-буровики, осуществляющие селекцию, анализ и разметку «сырых» данных со станций геолого-технологических исследований (ГТИ). Выбор способа разметки «сырых» данных для машинного обучения в интеллектуальных системах прогнозирования осложнений и аварийных ситуаций зависит от состава и качества данных ГТИ, периодичности их обновления и характера зависимости от геолого-геофизических характеристик породы и технологических процессов и др.

Большие высокочастотные геоданные — это наборы данных жизненного цикла мультисенсорного нефтегазового дела (со станций ГТИ — данные геолого-технологических исследований, данные ГИС — геофизических исследований скважин, ЭЦН — электроцентробежных насосов и других). Виды геоданных представлены на рисунке 1. Пример высокочастотных геоданных — это геоданные, которые записываются с частотой до 50 Гц на станциях геолого-технологических исследований при строительстве скважин. Сверхвысокочастотные геоданные с частотой дискретизации 100–1000 Гц генерируются в системах кибербезопасности и электропитания, используемых в нефтегазовом производстве.

Одна буровая установка генерирует от 10 до 150 Тб геоданных при работе на суше и на море соответственно [13]. С помощью технологий машинного обучения полученные большие геоданные можно использовать в интеллектуальных системах автоматизации и разработки нефтегазовых месторождений в целях повышения безопасности и снижения непроизводительного времени нефтегазового производства. На сегодняшний день основными высокочастотными данными, характеризующими скважинные процессы

в реальном масштабе времени, являются данные от станций геолого-технологических исследований, которые представляют собой временные ряды измеряемых параметров. Хотя большой интерес для повышения достоверности прогнозирования представляет использование геолого-геофизических характеристик, полученных при проведении поисковых изысканий и бурении разведочных скважин, а также использование измерений параметров бурового раствора при условии их проведения в масштабе времени близком к реальному.

Для машинного обучения в интеллектуальных комплексах прогнозирования осложнений используются архивные данные, полученные при строительстве нефтегазовых скважин в сходных геологических условиях. Первым подготовительным этапом является оценка состава и качества измерительных данных и их подготовка для формирования датасетов, используемых для машинного обучения. Формирование датасетов включает в себя проверку целостности данных, масштабирование, устранение аномальных выбросов и пропусков значений параметров и при необходимости их нормализацию [5]. Осложнения при бурении скважин могут отличаться по своей физической природе и характеристикам, в том числе в зависимости от проводимой в скважине операции. Для корректной работы модели прогнозирования осложнений целесообразно дополнительно реализовать возможность автоматического определения типов технологических операций по результатам обработки данных ГТИ [7].

Алгоритм определения основных технологических операций, выполняемых при строительстве скважин, реализован в виде скрипта на языке Python и выполняет анализ и обработку архивных данных ГТИ, представленных



Рис. 1. Виды геоданных: больше, высокочастотные и сверхвысокочастотные

Fig. 1. Types of geodata: big, high-frequency and ultra-high-frequency

the occurrence of complications during the construction of oil and gas wells.

Keywords

complications and emergency situations during well construction, drill string sticking, geological and technological measurement data, automated processing of field information, machine learning, neural network models, universal intelligent modules and complexes

в виде файлов в формате .las. Скрипт реализует несколько функций: очистка данных, удаление выбросов и определение типа технологических операций на основе обработки текущих параметров ГТИ.

Для работы алгоритма необходима установка утилит и библиотек:

- pandas: используется для манипуляции данными и анализа;
- lasio: позволяет читать и записывать файлы параметров ГТИ в формате .las, используемом для регистрации и хранения геофизических данных;
- datetime, timedelta: утилиты для управления датами и временем.

Скрипт осуществляет загрузку и преобразование .las файла (рис. 2):

- las.df преобразует данные .las файла в DataFrame pandas.
- pd.to_datetime: преобразует индекс DataFrame в формат даты и времени, где временные метки находятся в миллисекундах.

Скрипт определения межквартильного размаха (iqr) и удаления выбросов из DataFrame (рис. 3):

- quantile: рассчитывает первый и третий квартили;
- iqr: межквартильный размах, используемый для определения границ, что считается выбросом;
- dataframe_cleaned: DataFrame после

```
# Путь к файлу
file_path = r'путь_к_файлу'
# Чтение файла LAS
las = lasio.read(file_path)
df = las.df()
df.index = pd.to_datetime(df.index, unit='ms')
```

Рис. 2. Скрипт загрузки и преобразования .las файла

Fig. 2. Script for loading and converting .las file

```
import pandas as pd
def remove_outliers(dataframe):
    q1 = dataframe.quantile(0.25)
    q3 = dataframe.quantile(0.75)
    iqr = q3 - q1
    dataframe_cleaned = dataframe[(dataframe < (q1 - 1.5 * iqr)) | (dataframe > (q3 + 1.5 * iqr))]
    return dataframe_cleaned
```

Рис. 3. Скрипт определения межквартильного размаха и удаления выбросов

Fig. 3. Script for determining the interquartile range and removing outliers

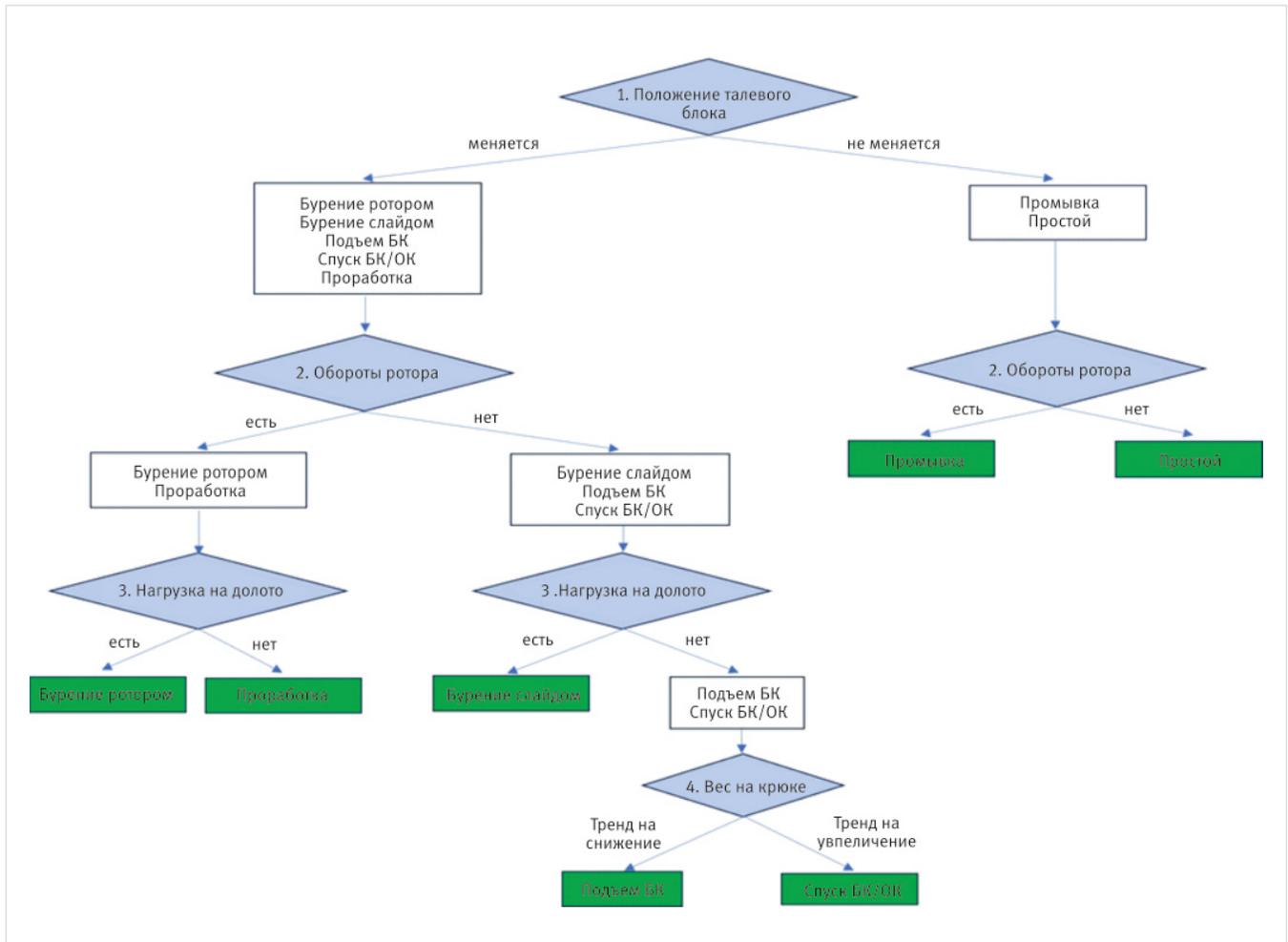


Рис. 4. Блок-схема алгоритма определения типа технологических операций
 Fig. 4. Block diagram of the algorithm for determining the type of technological operations

удаления выбросов.

Специально разработанный алгоритм определяет тип технологических операций, производимых в скважине. Алгоритм реализует анализ и интерпретацию различных параметров ГТИ (положение талевого блока, обороты ротора и нагрузка на долото). Блок-схема представлена на рисунке 4.

Пример работы алгоритма определения типа технологических операций представлен на рисунках 5, 6.

- `timeOfSticking`: конкретная временная метка, отмечающая конец временного окна для анализа.
- `operationType`: выполняет алгоритм для определения типа операции и выводит результат.

Зависимость возникновения осложнений от конструктивных особенностей скважины и используемого бурового оборудования можно учитывать включением в состав входных параметров моделей категориальных характеристик, влияющих на скважинные процессы [4]. Конкретный вариант разметки данных для прогнозирования осложнений и аварийных ситуаций при строительстве нефтяных и газовых скважин определяется структурой моделей нейросетей и методов машинного обучения, используемых в интеллектуальных системах. Ниже описаны варианты разметки данных ГТИ, применяемые для двух основных классов нейросетевых моделей.

Наиболее простой вариант разметки данных ГТИ используется для машинного

обучения одноклассовых нейросетевых моделей, реализующих бинарный прогноз развития ситуации: существует ли риск возникновения осложнения на определенном временном интервале или нет. При этом ключевое значение для корректной разметки архивных данных имеет правильное определение момента возникновения осложнения и временного интервала, включающего сценарий его развития.

Данные ГТИ представляют собой большие информационные массивы многомерных временных рядов значений разнородных параметров, зависящих от множества причин, в том числе от параметров управления — так называемый человеческий фактор, и не могут быть вручную проанализированы без априорной информации о проводимых операциях и режимах работы оборудования [6].

```
timeOfSticking = "2022-04-19 05:06:20"
operationType = algorithm(df_positive, timeOfSticking)
print(operationType)
```

Рис. 5. Скрипт запуска алгоритма определения типа технологических операций
 Fig. 5. Script for launching the algorithm for determining the type of technological operations

Для первоначальной локализации осложненных используются буровые журналы, отчеты об оптимизации бурения, акты расследования происшествий и др. документы, в которых приводится хронология технологических процессов, осуществляемых при строительстве скважины. При формализованном представлении информации в документации, например в формате WITSML, возможна программная реализация поиска, классификации и предварительной локализации осложнений, имевших место при строительстве скважин, что является актуальной задачей, учитывая большие объемы анализируемых архивных данных [10].

В буровых журналах часто встречаются ошибки в описаниях и приводимых характеристиках осложнений, а также неточности в определении хронологии их возникновения.

```
TIME
2022-04-19 00:20:20 4183.19 9.59 91.13 0.0 ... 1.39 0.0 56.10 0.0
2022-04-19 00:20:30 4183.19 10.55 90.45 0.0 ... 1.39 0.0 56.09 0.0
2022-04-19 00:20:40 4183.19 11.51 91.45 0.0 ... 1.39 0.0 56.01 0.0
2022-04-19 00:20:50 4183.19 12.51 90.80 0.0 ... 1.39 0.0 55.92 0.0
2022-04-19 00:21:00 4183.19 13.42 90.40 0.0 ... 1.39 0.0 55.91 0.0

[5 rows x 24 columns]
1. Положение талевого блока меняется
2. Есть обороты ротора
3. Нет нагрузки на долото
ПРОРАБОТКА
```

Рис. 6. Пример вывода результата работы алгоритма определения типа операции
 Fig. 6. Example of the output of the algorithm for determining the type of technological operations

Таблица 1. Пример формы разметки данных для одноклассовых моделей прогнозирования осложнений при строительстве нефтяных и газовых скважин

Table 1. Example of data labeling form for single-class models for predicting complications during construction of oil and gas wells

№ скв.	Категориальные параметры скважины		Тип операции	Тип осложнения	Время начала	Время конца	Глубина (м)	Диаграмма ГТИ
	Угол наклона, град.	Диаметр секции, мм						
241	25	311,2	Бурение ротором	Прихват	16.08.21 10.15	16.08.21 15.05	3 563–3 455	Ссылка
	25	311,2	Подъем КНБК	Прихват	17.08.23 12.15	17.08.21 18.05	4 825–4 770	Ссылка

Обязательным этапом является документальное подтверждение экспертами самого факта наступления осложнения, а также уточнение времени его возникновения и отклонений основных характеризующих параметров от их нормальных значений, которые должны проводиться по результатам анализа графического представления информативных параметров ГТИ.

Как показывает практика применения одноклассовых нейросетевых моделей прогнозирования осложнений и аварийных ситуаций при строительстве нефтяных и газовых скважин по результатам обработки реальновременных данных ГТИ, временной интервал прогноза при оптимальном

соотношении вероятности правильного предсказания и ложных тревог составляет порядка 10 минут [3]. При этом временной интервал разметки, включающий сценарий развития осложнения, можно ограничить значениями: за 30–60 минут до момента возникновения осложнения и 30 минут после, если предпринимались результативные действия по его ликвидации.

В таблице 1 приведен пример формы разметки данных для одноклассовых нейросетевых моделей для осложнений типа прихват, как самого распространенного и в то же время наиболее сложного для прогнозирования типа осложнений при строительстве скважин.

Вид диаграммы ГТИ, представляющей графическое представление параметров при осложнении типа прихват, приведен на рисунке 7.

Временной интервал прогноза возникновения осложнений при строительстве нефтяных и газовых скважин может быть увеличен при разработке и применении более сложных многоклассовых нейросетевых моделей [7]. Для синтеза их структуры необходимо выделить дополнительные классы событий, которые представляют собой различные отклонения от нормального режима бурения скважины и могут в конечном итоге привести к реальному возникновению осложнений. К таким событиям в случае прихватов можно,

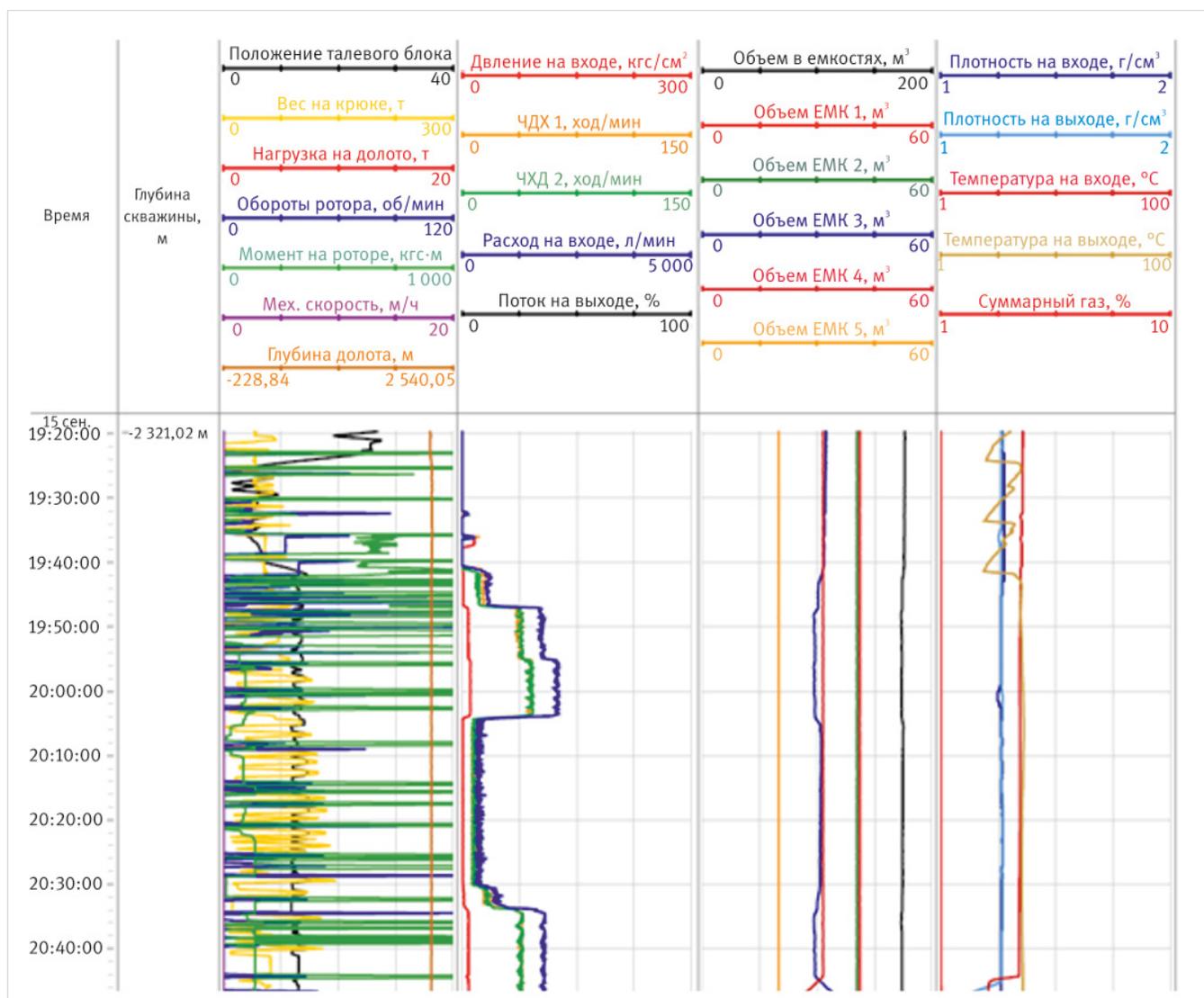


Рис. 7. Диаграммы параметров ГТИ при осложнении типа прихват
Fig. 7. Diagrams of GTI parameters for complications of the stuck

например, отнеси: затяжки, посадки, запорки ствола скважины и др. аномальные отклонения, характеризующиеся определенными параметрическими описаниями. Многоклассовые нейросетевые модели должны уметь идентифицировать и интерпретировать подобные события и их тренды развития во времени. Подобный подход лежит в основе машинного обучения робототехнических комплексов, обеспечивающих автоматическое бурение нефтегазовых скважин.

Однако не всякие возникающие отклонения от нормального режима при своевременном прогнозировании их возникновения следует рассматривать как несущие непосредственные риски возникновения осложнений, т. к. они могут быть обусловлены известными геолого-геофизическими факторами и особенностями технологического процесса бурения и могут быть нивелированы своевременными штатными действиями оператора буровой установки [11]. Для многоклассовых прогнозных моделей важным является выделение таких аномальных отклонений от нормы, которые несут в себе существенные риски развития и возникновения осложнений. Данные аномалии могут задаваться пороговыми значениями параметров, а также временными трендами развития этих отклонений.

Для машинного обучения многоклассовых нейросетевых моделей прогнозирования возникновения осложнений при строительстве нефтяных и газовых скважин необходимы разметки заданных классов отклонений от нормального режима бурения и степеней риска возникновения осложнений, с формированием предупреждающих сообщений оператору буровой установки и рекомендаций по принятию специальных мер для их устранения.

Пример разметки данных ГТИ для многоклассовой нейросетевой модели прогнозирования возникновения осложнений при строительстве нефтяных и газовых скважин представлен в таблице 2. В начальных столбцах формы разметки необходимо приводить категориальные параметры, которые задаются проектными решениями и не подлежат изменению оператором буровой установки, но в то же время могут оказывать влияние на возникновение осложнений. К таким параметрам могут относиться углы наклона скважины, диаметры секций, характеристики бурового оборудования и т. д. Для повышения достоверности прогнозов целесообразным является использование геолого-геофизических параметров, полученных по результатам предварительных исследований или разведочного бурения.

При разметке данных для многоклассовых нейросетевых моделей экспертами должны быть выделены временные интервалы, включающие участки с нормальным функционированием и отдельные сценарии возникновения осложнений, в которых должны быть детально размечены и описаны все отклонения от нормального режима работы в соответствии с введенной классификацией и тенденцией их развития. Важным этапом является подтверждение и уточнение времени наступления событий, которые осуществляются по результатам анализа графиков параметров ГТИ. В форме разметки данных должны быть отражены количественные характеристики отклонений основных параметров, характеризующих выделенную ситуацию. Эксперты оценивают степень риска возникновения осложнения по результатам

сравнения отклонений параметров с заданными пороговыми значениями, а также выявления и анализа трендов их развития во времени [12].

Заключение

Сложность при выполнении разметки данных и формировании датасетов для машинного обучения представляет то обстоятельство, что в процессе строительства нефтегазовых скважин и возникновения предпосылок к возникновению осложнений оператором буровой установки производятся активные действия по предупреждению развития негативных сценариев развития ситуаций. В форму разметки данных целесообразно включать описания действий буровика в конкретных ситуациях: это могут быть комментарии супервайзеров или оценки различных экспертов, в том числе проводящих разметку для машинного обучения. Практическое применение многоклассовых нейросетевых моделей для прогнозирования осложнений и аварийных ситуаций при строительстве нефтяных и газовых скважин по результатам обработки реальных данных ГТИ показывает, что временной интервал прогноза может составлять до 60 минут и более в зависимости от выявленных сценариев развития осложнений [7]. Таким образом, использование детальной разметки промышленных данных и многоклассовых нейросетевых моделей в интеллектуальных системах прогнозирования возникновения осложнений и аварийных ситуаций позволяет сделать значительный шаг не только в повышении безопасности и экономической эффективности производства, но и закладывает основу для автоматизации процессов строительства скважин и перехода к безлюдным роботизированным технологиям в нефтегазовой отрасли.

Итоги

Современное нефтегазовое производство генерирует большие потоки геоданных, восприятие и учет которых человеком невозможен без применения систем автоматизации с элементами искусственного интеллекта. Синтез структуры и эффективность применения систем искусственного интеллекта напрямую зависят от качества измерительных данных и их подготовки для формирования датасетов, используемых для машинного обучения. Разработанная методология позволяет автоматически осуществлять подготовку данных ГТИ и определение типов технологических операций. При этом применяемая технология детальной разметки данных, с выделением и идентификацией типовых отклонений от нормального режима работы, дает возможность синтезировать и применять многоклассовые нейросетевые модели прогнозирования возникновения осложнений. Исследования, проведенные в рамках и опытной эксплуатации экспериментального образца автоматизированной системы предупреждения осложнений, и аварийных ситуаций при строительстве скважин, показали, что реализация данного подхода позволяет значительно увеличить временной интервал и повысить достоверность прогноза возникновения аномальных ситуаций.

Выводы

Одним из мощнейших драйверов цифровой модернизации нефтегазовой отрасли является разработка и комплексное внедрение в производственные процессы

технологий искусственного интеллекта. В основе успешной реализации данного инновационного направления развития лежат организации работы с промышленными данными, привлечение экспертов, обучение и подготовка кадров. Подготовка и разметка данных для машинного обучения играют важную роль в определении структуры и обучении моделей нейросетей в интеллектуальных системах прогнозирования возникновения осложнений при строительстве нефтяных и газовых скважин. Данный подход имеет универсальный характер и может быть положен в основу автоматизации производственных процессов как при строительстве, так и при эксплуатации нефтяных и газовых скважин на базе внедрения интегрированных интеллектуальных систем разработки и эксплуатации нефтегазовых месторождений.

Литература

1. Дмитриевский А.Н., Дмитриев Е.С., Еремин Н.А. и др. Предупреждение осложнений в процессе бурения скважин методами машинного обучения // Технологии нефти и газа. 2024. № 3. С. 57–64.
2. Дмитриевский А.Н., Еремин Н.А., Черников А.Д. и др. Интеллектуальные системы предупреждения осложнений для безопасного строительства скважин // Безопасность труда в промышленности. 2022. № 6. С. 7–13.
3. Дмитриевский А.Н., Еремин Н.А., Черников А.Д. и др. Автоматизированная система предотвращения аварий при строительстве скважин // Нефтяное хозяйство. 2021. № 1. С. 72–76.
4. Черников А.Д., Еремин Н.А., Столяров В.Е. и др. Применение методов искусственного интеллекта для выявления и прогнозирования осложнений при строительстве нефтяных и газовых скважин: проблемы и основные направления решения. Георесурсы. 2020. Т. 22. № 3. С. 87–96.
5. Архипов А.И., Дмитриевский А.Н., Еремин Н.А. и др. Анализ качества данных станции геолого-технологических исследований при распознавании поглощений и газонефтеводопроявлений для повышения точности прогнозирования нейросетевых алгоритмов // Нефтяное хозяйство. 2020. № 8. С. 63–67.
6. Дмитриевский А.Н., Сбоев А.Г., Еремин Н.А. и др. Об увеличении продуктивного времени бурения нефтегазовых скважин с использованием методов машинного обучения. Георесурсы. 2020. Т. 22. № 4. С. 79–85.
7. Черников А.Д., Архипов А.И., Еремин Н.А. Расширение функционала программы интеллектуальной системы предупреждения осложнений и аварий при строительстве скважин (ИС ПОАС) // Современные технологии строительства и капитального ремонта скважин. Перспективные методы увеличения нефтеотдачи пластов. Сочи. 2023. Краснодар: ООО «Научно-производственная фирма «Нитпо», 2023. С. 19–25.
8. Aljurban M., Ramasamy J., Albassam M. et al. Deep learning and time-series analysis for the early detection of lost circulation incidents during drilling operations. IEEE Access, 2021, Vol. 9, P. 76833–76846. (In Eng).

Табл. 2. Пример разметки данных для многоклассовой нейросетевой модели прогнозирования возникновения осложнений
 Tab. 2. Example of data labeling for a multi-class neural network model for predicting the occurrence of complications

№ скв.	Диаметр секции, мм/дюймы	Тип операции	Режим работы, тип отклонения от нормального режима	Отклонения параметров от нормального режима			Время начала осложнения в соответствии с разметкой данных	Детальная разбивка операций внутри интервала времени: -6 ч +6 ч от начала осложнения		
				Вес на крюке	Давление на стояке	Крутящий момент		начало	конец	Описание работ
241	311,2/12,25	подъем КНБК	нормальный режим	-	-	-	-	03.05.2012 13:25	03.05.2012 15:05	подъем с обратной проработкой в интервале 4 680—4 580 м
241	311,2/12,25	подъем КНБК	запаковка	уменьшается	скачки +9—16 %	увеличивается	03.05.2012 15:10 03.05.2012 16:30 03.05.2012 17:25	03.05.2012 15:05	03.05.2012 17:50	подъем с обратной проработкой в интервале 4 580—4 530 м
241	311,2/12,25	подъем КНБК	запаковка	уменьшается	скачки +5—9 %	увеличивается	-	03.05.2012 17:50	03.05.2012 19:50	подъем с обратной проработкой в интервале 4 530—4 400 м
241	311,2/12,25	подъем КНБК	нормальный режим	-	-	-	-	03.05.2012 19:50	03.05.2012 21:40	подъем с обратной проработкой в интервале 4 400—4 300 м
241	311,2/12,25	подъем КНБК	запаковка	уменьшается	множественные скачки +5—10 %	увеличивается	03.05.2012 22:10 04.05.2012 00:40	03.05.2012 21:40	04.05.2012 01:15	подъем с обратной проработкой в интервале 4 300—4 210 м
241	311,2/12,25	подъем КНБК	нормальный режим	-	-	-	-	04.05.2012 01:15	04.05.2012 04:15	подъем с обратной проработкой в интервале 4 300—4 023 м
241	311,2/12,25	подъем КНБК	ТО	-	-	-	-	04.05.2012 04:15	04.05.2012 04:50	ремонт насоса
241	311,2/12,25	подъем КНБК	нормальный режим	-	-	-	-	04.05.2012 04:50	04.05.2012 05:40	подъем с обратной проработкой в интервале 4 023—3 980 м
241	311,2/12,25	подъем КНБК	запаковка	уменьшается	скачок +28 %	увеличивается	04.05.2012 05:40	04.05.2012 05:40	04.05.2012 07:50	подъем с обратной проработкой в интервале 3 980—3 960 м
241	311,2/12,25	подъем КНБК	затяжки, посадки	скачки +10—12 % Т скачки -18—20 % Т	нестабильное	увеличивается	04.05.2012 07:50 04.05.2012 08:55	04.05.2012 07:50	04.05.2012 10:00	расхаживание инструмента
241	311,2/12,25	подъем КНБК	запаковка	уменьшается	скачки +9—18 %	увеличивается	04.05.2012 10:45	04.05.2012 10:00	04.05.2012 12:50	подъем с обратной проработкой в интервале 3 960—3 920 м
241	311,2/12,25	подъем КНБК	затяжки, посадки, запаковка	скачки +10—12 % Т скачки -8—12 % Т	нестабильное скачки +7—13 %	увеличивается	04.05.2012 13:00 04.05.2012 14:40	04.05.2012 12:50	04.05.2012 15:15	подъем с обратной проработкой в интервале 3 920—3 900 м
241	311,2/12,25	подъем КНБК	прихват	скачки +40 Т скачки -40 Т	нестабильное	увеличивается	-	04.05.2012 15:15	04.05.2012 18:10	расхаживание инструмента
241	311,2/12,25	подъем КНБК	скачки давления	-	скачки 5 %	-	-	04.05.2012 18:10	04.05.2012 19:40	подъем с обратной проработкой в интервале 3 900—3 880 м
241	311,2/12,25	подъем КНБК	прихват	скачки +40 Т скачки -40 Т	нестабильное	увеличивается	04.05.2012 19:40	04.05.2012 19:40	04.05.2012 20:00	проработка интервалов затяжек, посадок
241	311,2/12,25	подъем КНБК	прихват	скачки +40 Т скачки -40 Т	нестабильное	увеличивается	04.05.2012 20:30	04.05.2012 20:00	04.05.2012 21:30	спуск КНБК с проработкой

9. Esmael B., Arnaout A.K., Fruhwirth R. et al. A statistical feature-based approach for operations recognition in drilling time series. *International journal of computer information systems and industrial management applications*, 2013, Vol. 5, P. 454–461. (In Eng).
10. Еремин Н.А., Чащина-Семенова О.К., Черников А.Д. и др. Экологичное бурение скважин с использованием интеллектуальной системы обработки высокочастотных данных со станций геолого-технологических измерений // Российская отраслевая энергетическая конференция. М.: Геомодель, 2023. С. 747–761.
11. Еремин Н.А., Дмитриевский А.Н., Черников А.Д. и др. Обеспечение экологической безопасности строительства газовых скважин на суше и на море на базе ИС ПОАС // Экологическая безопасность в газовой промышленности (ESGI-2023). М.: Газпром ВНИИГАЗ, 2023. 8 с.
12. Дмитриевский А.Н., Еремин Н.А., Гельфгат М.Я. и др. Обеспечение безопасности строительства скважин на основе использования интеллектуальных систем раннего предупреждения осложнений // Труды Российского государственного университета нефти и газа имени И.М. Губкина. 2022. № 1. С. 40–51.
13. Еремин Н.А. Черников А.Д., Столяров В.Е. Значение информации для цифровой трансформации при бурении и строительстве нефтегазовых скважин // Бурение и нефть. 2022. № 7–8. С. 8–18.

ENGLISH

Results

Modern oil and gas production generates large flows of geodata, the perception and accounting of which by humans is impossible without the use of automation systems with elements of artificial intelligence. The synthesis of the structure and the efficiency of the use of artificial intelligence systems directly depend on the quality of measurement data and their preparation for the formation of datasets used for machine learning. The developed methodology allows for the automatic preparation of GTI data and the definition of types of process operations. At the same time, the applied technology of detailed data markup, with the allocation and identification of typical deviations from the normal operating mode, makes it possible to synthesize and apply multi-class neural network models for predicting the occurrence of complications. Studies conducted within the framework and trial operation of an experimental sample of an automated system for preventing complications and emergency situations during well construction showed that the implementation of this approach can significantly increase the

time interval and improve the reliability of the forecast of the occurrence of abnormal situations.

Conclusions

One of the most powerful drivers of digital modernization of the oil and gas industry is the development and comprehensive implementation of artificial intelligence technologies in production processes. The successful implementation of this innovative development direction is based on the organization of work with field data, the involvement of the expert, and the education and training of personnel. The preparation and labeling of data for machine learning play an important role in determining the structure and training of neural network models in intelligent systems for predicting the occurrence of complications during the construction of oil and gas wells. This approach is universal and can be used as a basis for the automation of production processes both in the construction and operation of oil and gas wells based on the implementation of integrated intelligent systems for the development and operation of oil and gas fields.

References

- Dmitrievsky A.N., Dmitriev E.S., Eremin N.A. et al. Preventing complications during the drilling process wells using machine learning methods. *Science and technology of hydrocarbons*, 2024, issue 3, P. 57–54. (In Russ).
- Dmitrievsky A.N., Eremin N.A., Chernikov A.D. et al. Intelligent complication prevention systems for safe well construction. *Occupational safety in industry*, 2022, issue 6, P. 7–13. (In Russ).
- Dmitrievsky A.N., Eremin N.A., Chernikov A.D. et al. Automated system for preventing accidents during well construction. *Oil industry*, 2021, issue 1, P. 72–76. (In Russ).
- Chernikov A.D., Eremin N.A., Stolyarov V.E. et al. Application of artificial intelligence methods for identifying and predicting complications in the construction of oil and gas wells: problems and solutions. *Georesources*, 2020, Vol. 22, issue 3, P. 87–96. (In Russ).
- Arkhipov A.I., Dmitrievsky A.N., Eremin N.A., et al. Data quality analysis of the station of geological and technological researches in recognizing losses and kicks to improve the prediction accuracy of neural network algorithms. *Oil industry*, 2020, issue 8, P. 63–67. (In Russ).
- Dmitrievsky A.N., Sboev A.G., Eremin N.A. et al. On increasing the productive time of drilling oil and gas wells using machine learning methods. *Georesources*, 2020, Vol. 22, issue 4, P. 79–85. (In Russ).
- Chernikov A.D., Arkhipov A.I., Eremin N.A. Expanding the functionality of the program for the intelligent system for preventing complications and accidents during well construction (IS PCAW). *Modern technologies of construction and overhaul of wells. Promising methods for increasing oil recovery of formations*, Sochi, 2023. Krasnodar: «Research-and-Production firm «Nitpo» LLC, 2023. P. 19–25. (In Russ).
- Aljurban M., Ramasamy J., Albassam M. et al. Deep learning and time-series analysis for the early detection of lost circulation incidents during drilling operations. *IEEE Access*, 2021, Vol. 9, P. 76833–76846. (In Eng).
- Esmael B., Arnaout A.K., Fruhwirth R. et al. A statistical feature-based approach for operations recognition in drilling time series. *International journal of computer information systems and industrial management applications*, 2013, Vol. 5, P. 454–461. (In Eng).
- Eremin N.A., Chashchina-Semenova O.K., Chernikov A.D. et al. Eco-friendly drilling of wells using an intelligent system for processing high-frequency data from geological and technological measurement stations. *Russian Branch Energy Conference. Moscow: Geomodel*, 2023, P. 747–761. (In Russ).
- Eremin N.A., Dmitrievsky A.N., Chernikov A.D. et al. Ensuring the environmental safety of the construction of gas wells on land and at sea on the basis of IS POAS. *Environmental safety in the gas industry (ESGI-2023)*. Moscow: Gazprom VNIIGAZ, 2023, 8 p. (In Russ).
- Dmitrievsky A.N., Eremin N.A., Gelfgat M.Ya. et al. Ensuring safety of well construction based on intelligent early trouble prevention systems. *Proceedings of Gubkin University*, 2022, issue 1, P. 40–51. (In Russ).
- Eremin N.A., Chernikov A.D., Stolyarov V.E. The importance of information for digital transformation in drilling and construction of oil and gas wells. *Drilling and oil*, 2022, issue 7–8, P. 8–18. (In Russ).

ИНФОРМАЦИЯ ОБ АВТОРАХ | INFORMATION ABOUT THE AUTHORS

Еремин Николай Александрович, д.т.н., профессор, главный научный сотрудник, Институт проблем нефти и газа РАН, Москва, Россия

Eremin Nikolay Aleksandrovich, doctor of technical sciences, professor, chief researcher, Institute of oil and gas problems of the RAS, Moscow, Russia

Черников Александр Дмитриевич, к.т.н., ведущий научный сотрудник, Институт проблем нефти и газа РАН, Москва, Россия
Для контактов: cha60@mail.ru

Chernikov Alexander Dmitrievich, ph.d. of engineering sciences, leading researcher, Institute of oil and gas problems of the RAS, Moscow, Russia
Corresponding author: cha60@mail.ru