

## СЕКЦИЯ 7

# ИНФОРМАЦИОННЫЕ СИСТЕМЫ АНАЛИЗА ДАННЫХ

---

**В. В. ЕРЕМЕЕВ**

*Тюменский государственный университет, г. Тюмень*

**УДК 51-73, 550.83**

### ОБЗОР МЕТОДОВ АНАЛИЗА ДАННЫХ ДЛЯ ПРОГНОЗИРОВАНИЯ ЛИТОЛОГИИ В ПРОЦЕССЕ БУРЕНИЯ

**Аннотация.** Современные нефтяные и газовые месторождения требуют использования наклонно-направленного бурения для максимизации контакта скважин с целевыми пластами. Для этого используют геофизические приборы, при помощи которых определяют литологию породы. Однако существует проблема с зазором между приборами и долотом, который формирует «слепое окно» и может привести к неправильному позиционированию скважины. В статье рассмотрены несколько работ наиболее частым подходам для решения проблемы «слепого окна»: предсказание типа породы и предсказание показаний геофизических приборов на основе данных бурения.

**Ключевые слова:** литология в процессе бурения, генерация гамма каротажа, генерация плотности, корреляция между буровой информацией и литологией.

**Введение.** Современные нефтяные и газовые месторождения становятся все сложнее для экономически выгодной разработки. Компаниям приходится исследовать и разрабатывать тонкие пласты залежей углеводородов. Порой толщина этих пластов может быть не больше нескольких метров. Форма этих слоев сложнее, чем горизонтальная. Целью наклонно-направленного бурения является расположить скважины так чтобы она максимально контактировала с целевым пластом. Чтобы достичь этого скважина должна повторять изгибы пласта. Для этого инженеры по бурению используют геофизические приборы, отражающие физические свойства породы. Они позволяют понять находится ли скважина в целевом пласте или нет. Полученная информация с этих приборов позволяет инженерам вносить корректировки в направлении бурения. Существенным недостатком такого метода позиционирования является зазор между прибором и долотом, который бывает от 15 до 40 м. Такое «слепое окно» может привести к неправильному позиционированию скважины или даже к дорогостоящему повторному бурению. Пример такого позиционирования представлен на рис. 1.

**Основной раздел.** Для преодоления описанного выше слепого интервала исследователи выбрали два пути:

- по данным бурения предсказывать тип породы, где находится долото;
- по данным бурения предсказывать показания геофизических приборов в этой точке.

#### ***Работы по предсказанию типа породы.***

В работе [2] авторы непосредственно оценивали корреляционную взаимосвязь параметров бурения с литологией. Процесс бурения на долоте представляет собой сложный механический процесс в горной породе. Долото функционирует как вращающийся резец, управляемый несколькими параметрами, включая вес на долоте (WOB), размер долота и число оборотов

в минуту (RPM), которые влияют на скорость процесса прорезания породы. Авторы используют комбинацию этих параметров в виде "Силы прорезания  $F_c$  и изучают ее взаимосвязь со скоростью проходки (ROP). Для оценки взаимосвязи используют диаграммы рассеяния (рис. 2).

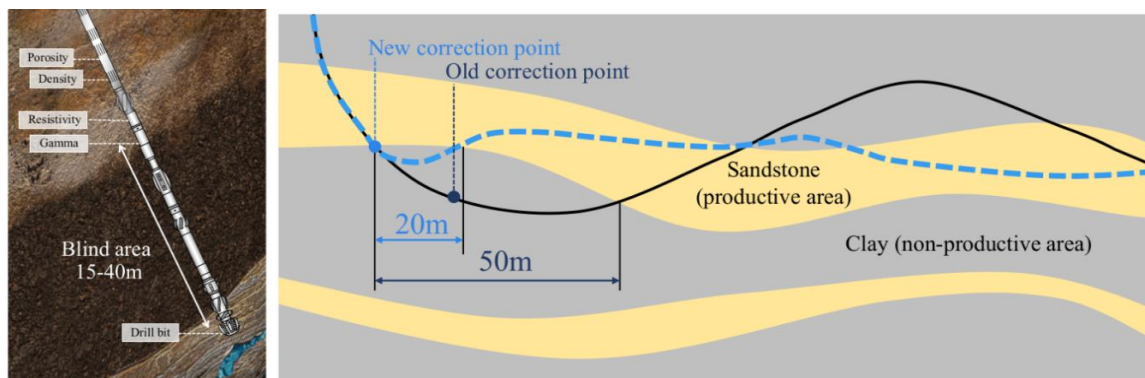


Рис. 1. Схематическое изображение буровой колонны (слева) и эффект своевременной коррекции траектории (справа):

черная кривая на рисунке показывает траекторию в случае, когда типы пород доступны только на расстоянии 15 м от бурового долота, синяя пунктирная кривая соответствует траектории, когда типы пород доступны на буровом долоте [1]

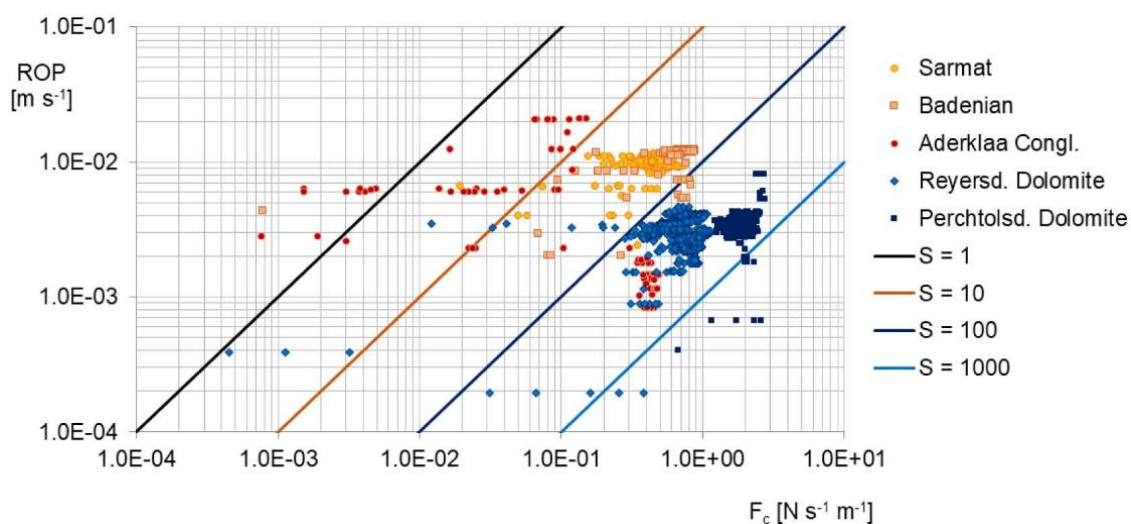


Рис. 2. Диаграмма рассеяния силы прорезания  $F_c$  к скорости проходки ROP для верхней части выбранного пласта одной скважины (цветом выделена литология; линией — параметр S)

Помимо изучения взаимосвязи авторы попытались разделить точки по литологии с использованием еще одного комбинированного параметра S, выражаемого формулой 1:

$$S = \frac{F_c}{ROP} \quad (1)$$

и исчисляется в МПа. На рис. 2 изображены линии, разделяющие по литологии разные точки. Они представляют отношение из формулы 1. В дальнейшем авторы работы исследовали взаимосвязь параметра S от глубины. На основе взаимосвязи S с глубиной сформировали еще два параметра: градиент параметра S от глубины и среднее значение S для разных пород, которыми им удалось разделить породы на «мягкие» и «твердые» (см. рис. 3).

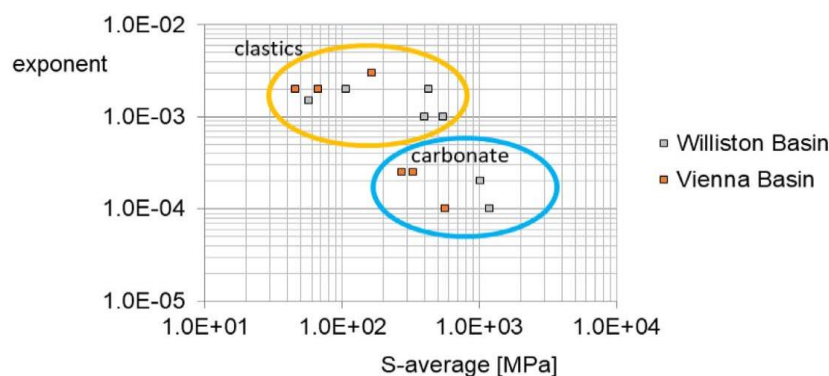


Рис. 3. Среднее значение S и градиент S в зависимости от глубины

В этой работе исследование проводилось на 2 скважинах. Тем не менее предложенные агрегированные параметры выделяют взаимосвязь параметров бурения от литологии.

Авторы следующей работы [3], также сосредоточились на прогнозировании литологии. В отличие от предыдущей работы авторы не стали создавать агрегирующие параметры, а использовали для предсказания следующие буровые параметры:

1. Torque — крутящий момент, расчетная величина, характеризующая действие приложенного усилия для создания вращения долота;
2. Stratesum — общая скорость хода насоса;
3. Bit\_RPM — обороты в минуту на долоте;
4. Pump — давление насоса;
5. FlowOut — расход на выходе из скважины;
6. ROP\_avg — средняя скорость проходки;
7. TOTGAS — общее содержание газа в буровом растворе;
8. WOB — вес буровой колонны на долоте;
9. ECDBIT — плотность бурового раствора.

В качестве предсказательной модели авторы исптали алгоритм случайного леса, градиентный бустинг над линейной моделью, метод опорных векторов, метод ближайших соседей. В итоге они остановились на использовании алгоритма случайного леса для эффективного разделения глинистых пород, мергелей и песчаников. В результате они достигли уровня точности в 96%. Каротажи, с входными данными представлены на рис. 4.

Для обучения и оценки качества использовали записи с 2 скважин, что кажется должно быть недостаточно для оценки качества модели. Тем не менее, работа указывает на возможность применения моделей машинного обучения для предсказания литологии по данным бурения.

В третьей работе о литологии [1] исследование проводили на 27 скважинах в Западной Сибири. Авторы очень хорошо описали проблематику для чего надо определять литологию на долоте в процессе бурения. Еще ярко описана сложность в сборе данных для анализа, так как изначально запись параметров осуществлялась для контроля процесса бурения, а не для анализа и поиска зависимостей. К примеру, данные для разных групп параметров хранятся в разных файлах, большинство записей привязано к временной шкале, но есть записи, привязанные к глубинной шкале. В большей части скважин есть дисбаланс классов в пользу песчаников. В связи с этим минорные классы глины и твердые породы объединили в одну группу. В качестве метрики авторы использовали RocAUC, PR AUC и специфичную точность, вычисленную от

подсчета правильно классифицированных интервалов. В работе подробно описано как создавали новые атрибуты свойственные временным рядам, такие как: сглаженные производные, отложенные на разных интервалах значения исходных параметров, стандартное отклонение внутри интервала 0.1 метра и отложенная информация с геофизических приборов, по которым традиционно размечается литология. Исследователи отбирали параметры следующим способом: сначала обучали модель градиентного бустинга на деревьях для разных комбинаций параметров, а после замеряли метрики на отдельной скважине вне обучающей выборки. В результате были отобраны 11 атрибутов. После апробированы логистическая регрессия, бустинг на деревьях и полно-связная нейронная сеть. В итоге по балансу метрик авторы выбрали градиентный бустинг. Результаты работы модели на нескольких скважинах представлены на рис. 5.

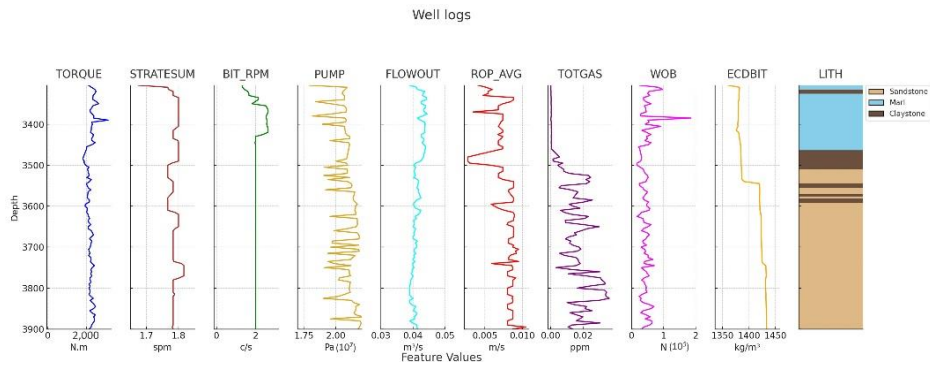


Рис. 4. Каротаж с основными характеристиками, отобранными для исследования

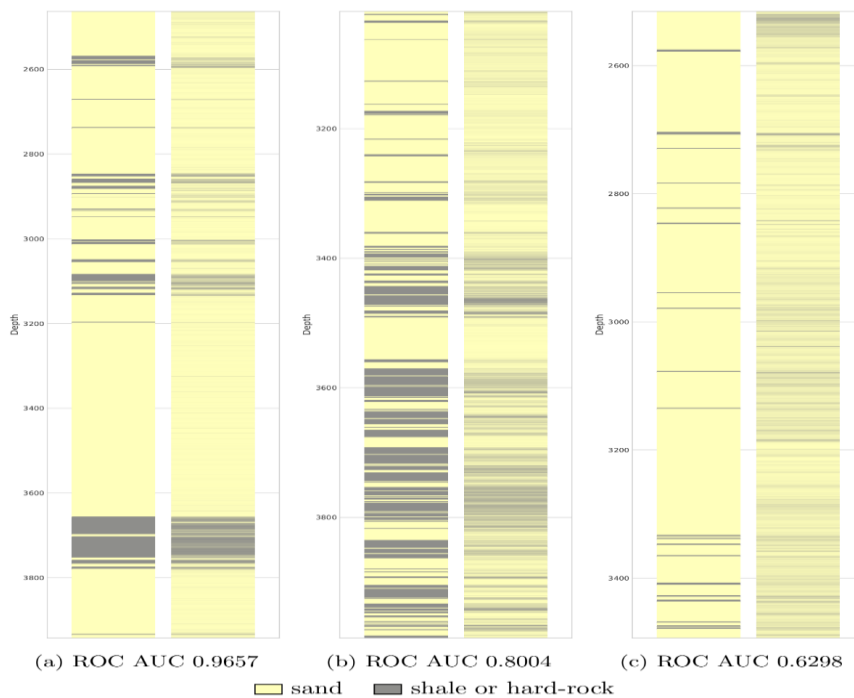


Рис. 5. Примеры классификации литотипов для трех скважин с разным достигнутым качеством: от одного из лучших слева через среднее в середине до одного из худших справа: в каждом разделе в левой колонке показаны истинные значения литотипов: желтым цветом обозначены пески, серым — сланцы и твердые породы; в правой колонке — соответствующие вероятности литотипов, выданные классификатором

Возможность тестирования модели на отдельной скважине позволяет избежать утечки информации о целевом показателе. Если в общем оценивать это направление, то эффективность обучения моделей очень сильно зависит от качества разметки литологии интерпретаторами ГИС и количества размеченных данных.

**Работы по предсказанию показания геофизических приборов.** Следующий подход для устранения проблемы это прогнозирование показаний геофизических приборов. Первая из исследованных работ была работа [4]. Исследование проводилось на одной скважине в бассейне реки Нигер. Авторы в своем исследовании сосредоточились на проверке эффективности нескольких алгоритмов машинного обучения, в частности тех, которые имеют возможность моделировать последовательность. Входные параметры бурения агрегированы в один параметр нормализованная по глубине гидромеханическая удельная энергия:

$$HMSE_{dn} = \left( \frac{WOB}{A} + \frac{120\pi RPM \cdot TRQ}{A \cdot ROP} + \frac{1154\eta \Delta P_b Q}{A \cdot ROP} \right) \phi_0 e^{-KZ}, \quad (2)$$

где WOB — вес буровой колонны на долоте; Ab — площадь долота; N — скорость вращения; ROP — скорость проходки; T — крутящий момент;  $\eta$  — коэффициент снижения гидравлической энергии;  $\Delta P_b$  — перепад давления на форсунке; Q — расход бурового раствора;  $\phi_0$  — поверхностная пористость; Z — истинная глубина; K — коэффициент уплотнения.

Для оценки связи  $HMSE_{dn}$  и гамма каротажа авторы выровняли по глубине оба параметра, и построили по ним график от глубины (рис. 6) и диаграмму рассеяния (рис. 7).

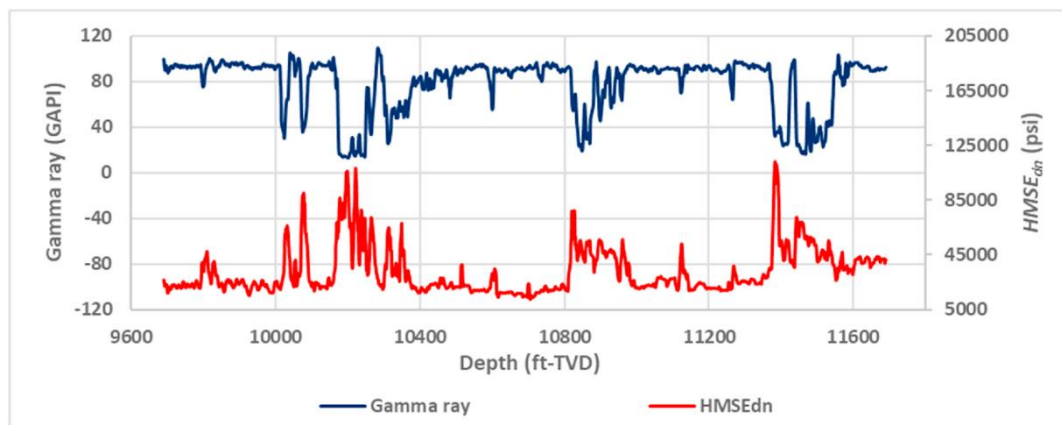


Рис. 6. График ГК и  $HMSE_{dn}$  от глубины

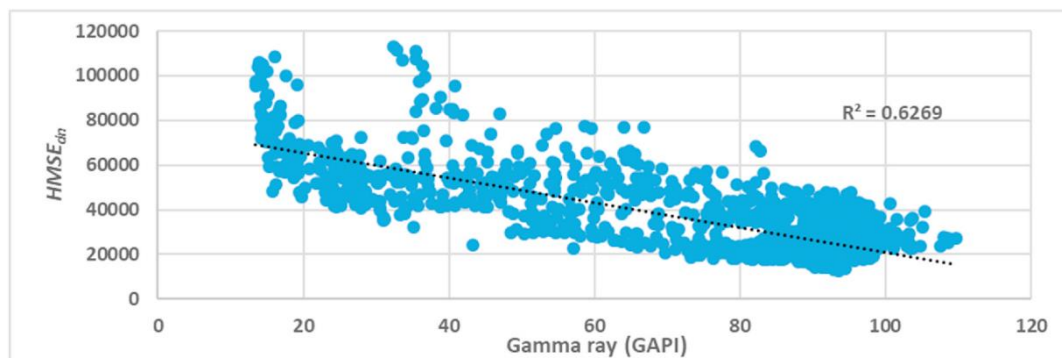


Рис. 7. Диаграмма рассеяния ГК и  $HMSE_{dn}$

Глядя на оба графика, можно заметить обратную корреляционную взаимосвязь. На основе этой взаимосвязи, авторы протестировали несколько алгоритмов машинного обучения для генерации ГК:

- 1) полносвязную нейронную сеть;
- 2) простую рекуррентную нейронную сеть RNN;
- 3) модель нелинейной авторегрессии с экзогенными входами NARX;
- 4) сети долгой краткосрочной памяти LSTM-RNN;
- 5) управляемый рекуррентный блок GRU;
- 6) временные сверточные сети TCN.

По результатам тестирования простая рекуррентная сеть показала лучшие результаты. Эта работа показала наличие взаимосвязи между параметрами бурения и гамма-каротажем.

Схожим подходом воспользовались и другие ученые из Австралии в своей работе [5], но они генерировали показания гамма-гамма плотностного каротажа. Основной целью предсказания плотности на основе других атрибутов является снижение затрат на использование дорогостоящего оборудования. В качестве исходных данных информация с 6 скважин на метан-угольных пластах, залегающих на глубине до 1000 метров. Авторы во входные параметры для предсказания включили гамма каротаж, на основании взаимосвязи гамма-каротажа и плотности породы. Диаграмма рассеяния представлена на рис. 8.

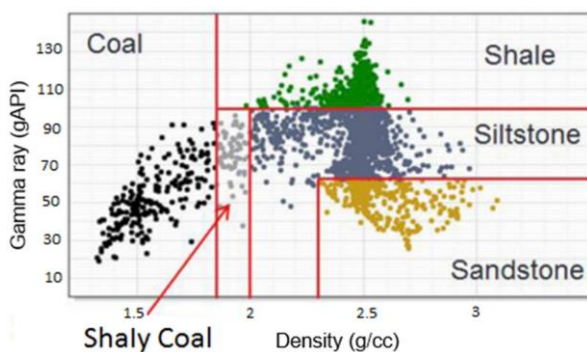


Рис. 8. Диаграмма рассеяния для плотности и ГК с разметкой литологии с 1200 скважин Бассейна Сюрат в Индонезии [6]

В итоге обучив модель на основе алгоритма XGBoost, получились результаты генерации модели на рис. 9.

Можно сказать псевдоплотностной каротаж отражает продолжительные изменения плотности, хоть и имеет тенденцию к среднему значению. Важность признаков, выделенных моделью XGBoost представлена на рис. 10.

Сразу видно, что глубина оказывается наибольшее влияние на итоговое значение. Этому есть разумная интерпретация, все скважины расположены не далеко друг от друга, поэтому литология у скважин не сильно отличается друг от друга. Второй по важности параметр, гамма каротаж, также отражает литологию. Получается, что модель применима только в границах этого месторождения.

Направление предсказания показаний геофизических приборов выглядит более перспективным, так как в процессе бурения скважин накапливается большой набор целевых показателей без дополнительной работы по интерпретации.



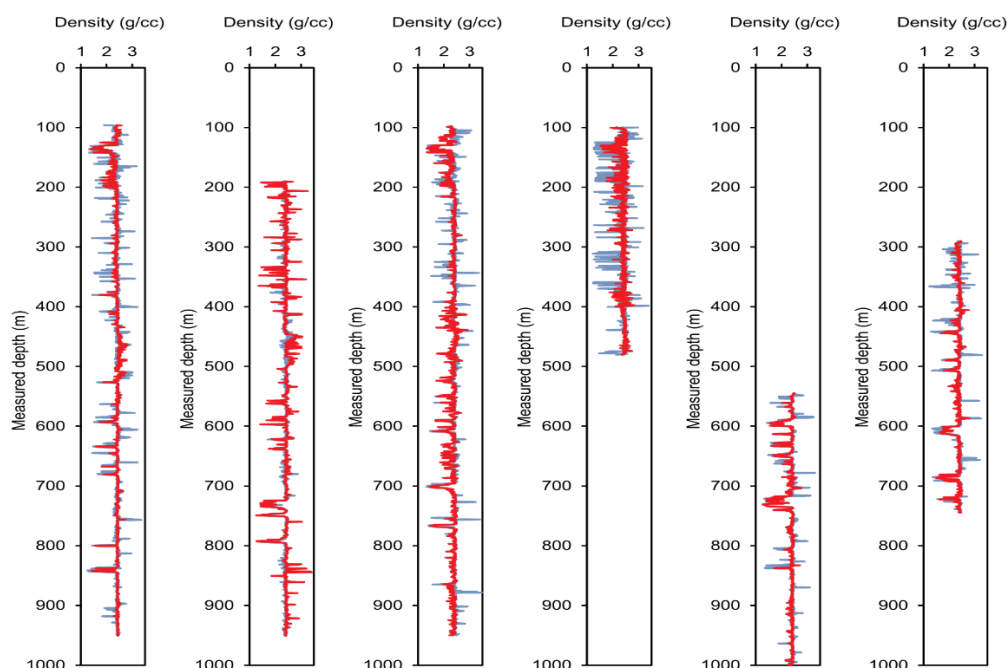


Рис. 9. Результаты сгенерированных псевдоплотностных каротажей для 6 скважин: синие линии — исходные каротажи плотности, красные линии — псевдоплотностные каротажи, генерированные с помощью XGBoost

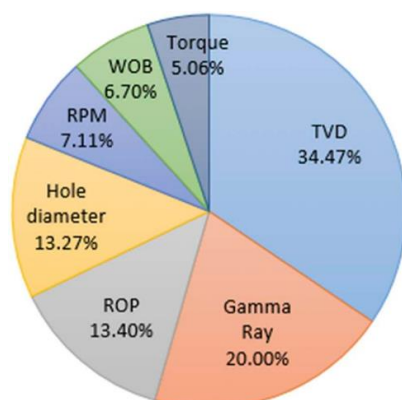


Рис. 10. Важность признаков, выделенных моделью XGBoost

**Заключение.** Для решения проблемы «слепого окна» в процессе бурения есть два популярных подхода: предсказание литологии на долоте и опережающее предсказание показания геофизических приборов. У каждого из подходов есть свои минусы. К недостаткам первого подхода можно отнести небольшой набор интерпретированных данных, предвзятость при разметке в пользу индивидуального опыта интерпретатора. К недостаткам второго подхода можно отнести не линейную взаимосвязь параметров бурения от геофизических свойств породы (твердость породы не всегда соотносится с ее естественной радиоактивностью), наличие ошибок записях параметров во время бурения.

Чтобы устранить недостатки первого подхода потребуются крупные финансовые расходы на повторную перекрестную разметку большого набора данных группами интерпретаторов. Для второго подхода — нужно будет найти способ устранить ошибки в записях, подобрать алгоритм для улавливания взаимосвязи параметров бурения со свойствами наиболее

важных для наклонно-направленного бурения пород. В связи с вышеизложенным, в первую очередь надо сосредотачиваться на генерации показаний геофизических приборов по параметрам бурения, а если быть точнее на генерации гамма-каротажа.

#### СПИСОК ЛИТЕРАТУРЫ

1. Klyuchnikov N. Data-driven model for the identification of the rock type at a drilling bit / N. Klyuchnikov, A. Zaytsev, A. Gruzdev [et al.] — DOI 10.1016/j.petrol.2019.03.041. — Direct text // Journal of Petroleum Science and Engineering. — July 2019. — Vol. 178. — Pp. 506-516.
2. Lamik-Thonhauser B. Correlation Between Drilling Parameters and Lithology — The Hidden Geological Information of Drilling Data / B. Lamik-Thonhauser, J. Schoen, C. Koller, A. Arnaout. — DOI 10.2118/192916-MS. — Direct text // International Petroleum Exhibition & Conference — Abu Dhabi, UAE, November 2018.
3. Khalifa H. Machine Learning-Based Real-Time Prediction of Formation Lithology and Tops Using Drilling Parameters with a Web App Integration / H. Khalifa, O.S. Tomomewo, U.F. Ndulue, B.E. Berrehal. — DOI 10.3390/eng4030139 — Direct text // Eng-Advances in Engineering 4 (Artificial Intelligence and Data Science for Engineering Improvements). — 2023. — № 3. — Pp. 2443-2467.
4. Osarogiagbon A. Gamma ray log generation from drilling parameters using deep learning / A. Osarogiagbon, O. Oloruntobi, F. Khan, R. Venkatesan, S. Butt. — DOI 10.1016/j.petrol.2020.107906. — Direct text // Journal of Petroleum Science and Engineering. — 2020. — Vol. 195. — Pp. 107906.
5. Zhong R. Generating pseudo density log from drilling and logging-while-drilling data using extreme gradient boosting (XGBoost) / R. Zhong, R. Johnson, Z. Chen. — DOI 10.1016/j.coal.2020.103416. — Direct text // International Journal of Coal Geology. — 2020. — Vol. 220. — Pp. 103416.
6. Zhang M. CBM fluvial depositional facies modelling and static upscaling workflow: a case study based on integration of cores, logs and deposition environment in Surat Basin / M. Zhang, G. Thomas, Y. Yang [et al.]. — DOI 10.2118/186340-MS. — Direct text // SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Jakarta, Indonesia, October 2017.