

СБОРНИК ТЕЗИСОВ ДОКЛАДОВ

ТОМ 2

XII Международная научная конференция

МОЛОДЫЕ – НАУКАМ О ЗЕМЛЕ

30 – 31

марта 2026 г.



УДК 55 : [550.812+622] : 001.895 + [622.276+622.279] (082)
ББК 26.347

Молодые – наукам о Земле: в 4 т. Материалы XII Международной научной конференции молодых ученых «Молодые – наукам о Земле» - М.: Издательство МГРИ, 2026.

Т. 2: ИННОВАЦИОННЫЕ ТЕХНОЛОГИИ ГЕОЛОГИЧЕСКОЙ РАЗВЕДКИ И ДЕЛА, НЕФТЕГАЗОВОЕ ДЕЛО / ред. коллегия: Ю.П. Панов, Г.А. Мамед-Заде - М.: Издательство МГРИ, 2026 – 570 с.

ISBN 978-5-907595-25-5
УДК 55 : [550.812+622] : 001.895 + [622.276+622.279] (082)
ББК 26.347

ISBN 978-5-907595-25-5 (Том 2)

© РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГЕОЛОГОРАЗВЕДОЧНЫЙ
УНИВЕРСИТЕТ ИМЕНИ СЕРГО ОРДЖОНИКИДЗЕ, 2026

Прогнозирование интервалов продуктивного пласта с использованием алгоритма XGBoost на основе данных геофизических исследований скважин

***Буй Т.Н.* (Ханойский университет горного дела и геологии, buithingan@humg.edu.vn),
Лэ Н.А. (Ханойский университет горного дела и геологии, lengocanh@humg.edu.vn),
Нгуэн Д.М. (Ханойский университет горного дела и геологии,
nguyenduymuoi@humg.edu.vn)***

Аннотация

Определение интервалов продуктивного пласта является важным этапом при оценке потенциала коллектора и подсчете запасов углеводородов. Разграничение продуктивных и непродуктивных интервалов позволяет геологам-нефтяникам принимать обоснованные решения относительно размещения скважин, выбора объектов бурения, а также применения эффективных методов управления и разработки месторождения.

Традиционные методы определения интервалов продуктивного пласта обычно основаны на использовании пороговых значений (cut-off values) некоторых петрофизических параметров, таких как глинистость, пористость, проницаемость и насыщенность флюидами. Однако данный подход имеет ряд ограничений, включая возможное накопление ошибок при интерпретации данных геофизических исследований скважин (ГИС), а также значительные временные затраты при выполнении ручного анализа и интерпретации данных.

Применение моделей машинного обучения, использующих данные ГИС, открывает возможности для автоматизации процесса интерпретации, повышения скорости обработки данных и обеспечения большей согласованности при выделении интервалов продуктивного пласта. В данной работе алгоритм Extreme Gradient Boosting (XGBoost) применяется для классификации интервалов продуктивного пласта на основе данных ГИС месторождения Y, расположенного в бассейне Кыулунг (Вьетнам).

Ключевые слова

Бассейн Кыулунг (Вьетнам), геофизические исследования скважин, интервалы продуктивного пласта, интерпретация, XGBoost

Теория

Интервалы продуктивного пласта представляют собой участки пород-коллекторов, насыщенные нефтью, газом или конденсатом и способные отдавать флюиды при разработке. Их выделение часто основывается на данных ГИС по пороговым значениям пористости, проницаемости и насыщенности [1].

Эта концепция включает не только технические ограничения, но и коммерческие аспекты разработки месторождения. Она учитывает такие факторы, как качество коллектора, характеристики пластовых флюидов и экономические условия, с целью определения интервалов, потенциально пригодных для промышленной добычи. Даже для традиционных нефтегазовых месторождений с длительной историей изучения до настоящего времени не существует унифицированного метода определения интервалов продуктивного пласта [3].

В связи с этим пороговые значения, используемые для выделения таких интервалов, носят в определенной степени субъективный характер и могут различаться в зависимости от подходов нефтегазовых компаний, а также требований регулирующих органов. Таким образом, для разных геологических и литологических условий могут применяться разные наборы пороговых значений при оценке интервалов продуктивного пласта. Типичные пороговые значения, предложенные для песчаников и карбонатов, показаны в таблице 1 [2].

Таблица 1

Пороговые значения при оценке интервалов продуктивного пласта для песчаников и карбонатов

Литология	Параметры	Диапазон значений
Песчаники	V_{sh}	0.3-0.5
	ϕ	0.06-0.08
	S_w	0.5-0.6
Карбонаты	V_{sh}	0.3-0.5
	ϕ	0.04-0.05
	S_w	0.5-0.6

Алгоритмы машинного обучения могут эффективно применяться для автоматического определения интервалов продуктивного пласта на основе больших массивов данных, что позволяет минимизировать ошибки и сократить время анализа и интерпретации.

Целью данной работы является оценка применимости модели XGBoost для автоматического определения интервалов продуктивного пласта в районе месторождения Y, расположенного в бассейне Кыулунг (Вьетнам).

Данные, использованные в этом исследовании, включают данные ГИС трех скважин А, В и С на месторождении Y. Рассматриваемые глубинные интервалы в этих скважинах соответствуют отложениям позднего олигоцена. Входными каротажными кривыми для модели являются DT, GR, NPHI, RHOB и LLd, а целевой переменной для прогнозирования служит кривая PAYFLAG. Кривая PAYFLAG формируется на основе пороговых значений петрофизических параметров в сочетании с экспертной геологической интерпретацией и с учетом региональных литолого-фациальных особенностей осадочных пород (Рис. 1).

Результатом работы модели является классификация каждой точки наблюдения на заданной глубине на дискретные категории, например, {непродуктивная|продуктивная}. При этом каждая точка наблюдения соответствует вектор признаков $x \in R^p$ и дискретную метку $y \in \{0,1\}$, где значение 0 соответствует непродуктивному интервалу, а значение 1 – продуктивному интервалу.

В типичном процессе машинного обучения, данные из скважины А используются в качестве обучающего набора, данные из скважины В – в качестве проверочного набора для выбора гиперпараметров с целью определения оптимальной модели, а данные из скважины С – в качестве тестового набора для оценки применимости модели к скважине, которая не использовалась в процессе обучения [4].

Для оценки эффективности модели в задаче классификации используются метрики, включая точность (Precision) и полнота (Recall), которые рассчитываются по следующим формулам:

$$\text{Точность} = \frac{\text{ИП}}{\text{ИП} + \text{ЛП}} \quad (1)$$

$$\text{Полнота} = \frac{\text{ИП}}{\text{ИП} + \text{ЛО}} \quad (2)$$

В которых:

ИП – Истинно-положительный результат (ИП – True Positives) - модель правильно определила Класс 1 (продуктивный интервал), ИО – Истинно-отрицательный результат (ИО, True Negative): модель правильно определила Класс 0 (непродуктивный интервал), ЛП – Ложно-положительный результат (ЛП, False Positives): модель неправильно определила класс

1, ЛО – Ложно-отрицательный результат (ЛО, False Negatives): модель неправильно определила класс 0.

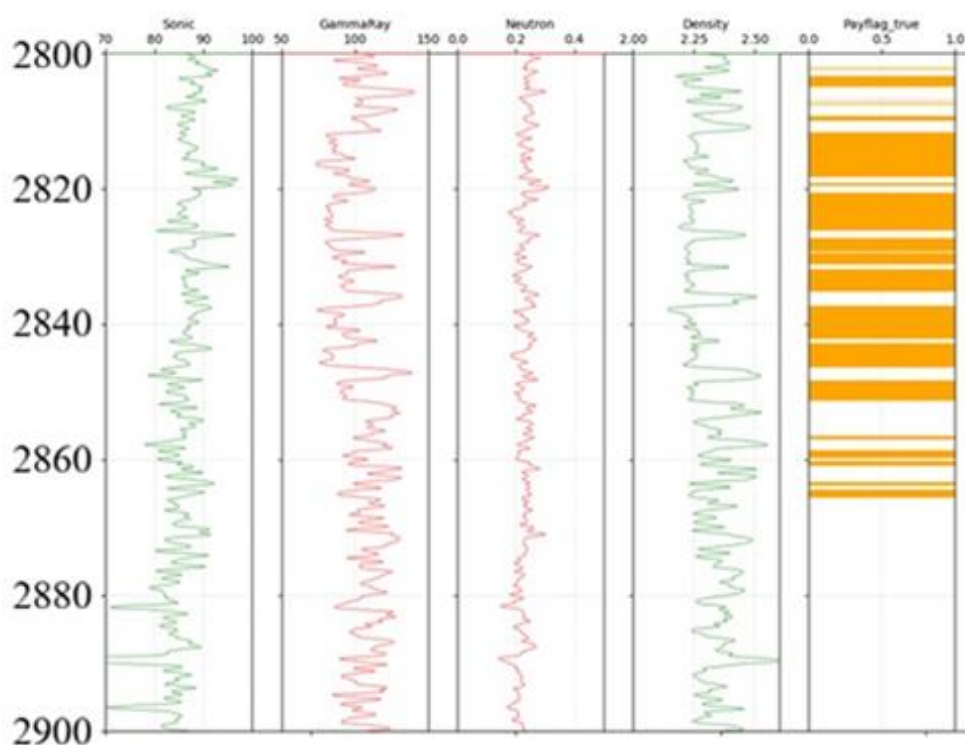


Рисунок 1. Примеры некоторых кривых геофизических исследований скважин, использованных в данном исследовании

Матрица ошибок (Confusion matrix) используется для наглядного представления количества точек данных, соответствующих результатам классификации при прогнозировании интервалов продуктивного пласта в скважине С (Рис. 2).

Из матрицы видно, что 5546 точек измерений были правильно классифицированы как непродуктивные (истинно-отрицательные), 422 точки были правильно классифицированы как продуктивные (истинно-положительные), 32 точки были ошибочно классифицированы как продуктивные (ложно-положительные), и 123 точки были ошибочно классифицированы как непродуктивные (ложно-отрицательные).

Точность (Precision) = 93% означает, что доля правильно предсказанных продуктивных интервалов среди всех точек наблюдений, классифицированных моделью как продуктивные, составляет 93%.

Полнота (Recall) = 77%, что означает, что модель смогла обнаружить 77% фактических интервалов продуктивного пласта, в то время как 23% продуктивных интервалов были ошибочно классифицированы как непродуктивные. С практической точки зрения это эквивалентно потере части чистой толщины продуктивного пласта при прогнозировании.

Относительно низкое значение показателя полноты (Recall) часто наблюдается в задачах классификации продуктивных интервалов пластов, поскольку данные по продуктивным и непродуктивным интервалам обычно являются несбалансированными. Как правило, количество непродуктивных интервалов значительно превышает количество продуктивных, вследствие чего модель может отдавать предпочтение доминирующему классу. Несмотря на это, полученные значения метрик свидетельствуют о высокой способности модели XGBoost выявлять продуктивные интервалы по данным каротажа.

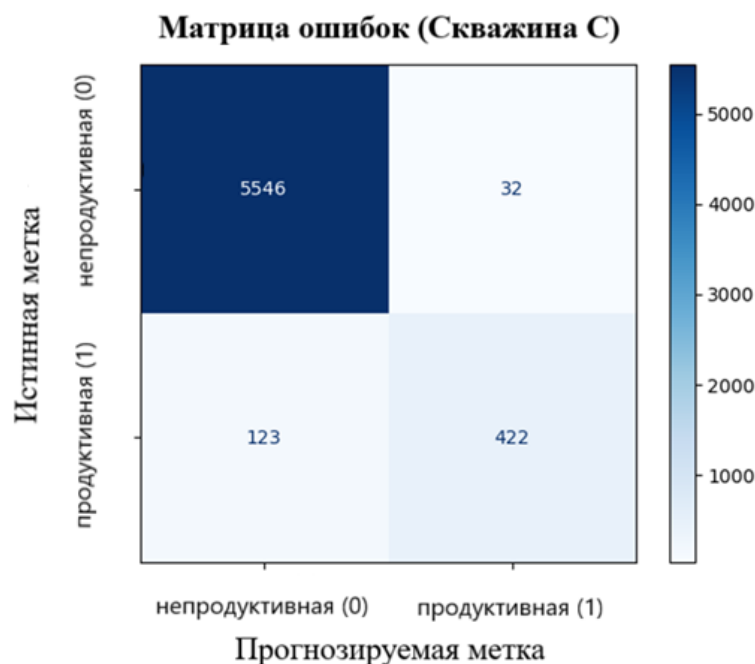


Рисунок 2. Матрица ошибок, оценивающая модель прогнозирования продуктивных интервалов пласта скважины С

Выводы

Результаты исследований показывают, что модель XGBoost является и надежным инструментом для классификации продуктивных интервалов пласта на основе данных каротажа скважин. Однако для повышения точности модели и улучшения ее способности выявлять тонкие пласты-коллекторы необходимы более глубокая оптимизация гиперпараметров, а также использование более сложных алгоритмов машинного обучения.

Библиография

1. Cheatwood C.J. and Guzman A.E. Comparison of Reservoir Properties and Development History: Spraberry Trend Field, West Texas, and Chicotepec Field, Mexico // SPE International Petroleum Conference and Exhibition. – 2002. – DOI: 10.2118/74407-MS
2. Worthington, P.F., and Cosentino L. The Role of Cut-offs in Integrated Reservoir Studies // SPE Reservoir Evaluation & Engineering. – 2005. – Vol 8(4). – P. 276-290. – DOI: <https://doi.org/10.2118/84387-PA>
3. Worthington P.F., Majid A.A. The Use of Net-Pay Concepts in the exploitation of shale-gas deposits // Journal of Petroleum Science and Engineering. – Vol. 120. – 2014. – P. 78-85. – DOI: <https://doi.org/10.1016/j.petrol.2014.05.014>
4. Saroji S., Winata E., Hidayat P.P.W., Prakoso S. and Herdiansyah F. The implementation of machine learning in lithofacies classification using multi well logs data // Aceh Int J Sci Technol. – 2021. – P. 9-17.