

# Разработка методики прогнозирования пластовых свойств нефти при помощи машинного обучения

Фрейман О.А.<sup>1</sup>, Еремин Н.А.<sup>1,2</sup>

<sup>1</sup>САФУ имени М.В. Ломоносова, Архангельск, Россия, <sup>2</sup>Институт проблем нефти и газа РАН, Москва, Россия  
frejman.o@edu.narfu.ru

## Аннотация

Пластовые свойства нефти являются одними из основных параметров для решения многочисленных проблем разработки залежей. Экспериментальное определение пластовых свойств нефти требует много времени и затрат. Разработка точной и эффективной модели прогнозирования является актуальной проблемой. Анализ данных пластовых свойств нефти необходим для обоснования эффективного регулирования разработки месторождения. В последние десятилетия были разработаны различные модели для оценки коллекторских свойств пластовых флюидов. Эти ранее опубликованные модели являются эмпирическими, композиционными и основанными на нейронных сетях. В данном исследовании использован метод машинного обучения для прогнозирования показателей разработки нефтяного месторождения и расчета пластовых свойств нефти.

## Материалы и методы

Свойства пластовых флюидов были взяты из открытой базы данных по месторождению Вольве, Северное море, норвежский сектор. Методы машинного обучения легли в основу определения пластовых свойств флюидов и расчета технологических показателей разработки.

## Ключевые слова

пластовые свойства, нефть и газ, машинное обучение, искусственный интеллект

Работа проведена в рамках плана научно-исследовательских работ очной аспирантуры в Северном (Арктическом) федеральном университете и выполнения государственного задания, номер гос. рег. № НИОКТР в РОСРИД 122022800270-0.

## Для цитирования

Фрейман О.А., Еремин Н.А. Разработка методики прогнозирования пластовых свойств нефти при помощи машинного обучения // Экспозиция Нефть Газ. 2023. № 7. С. 118–120. DOI: 10.24412/2076-6785-2023-7-118-120

Поступила в редакцию: 17.10.2023

## AUTOMATION

UDC 622.276 | Original Paper

## Development of a methodology for predicting reservoir properties of oil using machine learning methods

Freiman O.A.<sup>1</sup>, Eremin N.A.<sup>1,2</sup>

<sup>1</sup>Northern (Arctic) Federal University, Arkhangelsk, Russia, <sup>2</sup>Oil and Gas Research Institute RAS, Moscow, Russia  
frejman.o@edu.narfu.ru

## Abstract

Reservoir properties of oil are necessary to justify the effective regulation of field development. Measurement accuracy in field development depends on reservoir data (eg material balance calculations, reserves estimation, predictive data analysis). Incorrect measurement of reservoir properties can lead to serious errors in the calculation results. In the literature, the influence of reservoir data uncertainty on test results was considered, for example, in material balance equations and estimates of hydrocarbon reserves and the release of more volatile fluids. In recent decades, various models have been developed to assess the reservoir properties of formation fluids, such as empirical, compositional and based on neural networks. In this study, a machine learning method will be used to predict the performance indicators of an oil field and calculate reservoir fluid properties.

## Materials and methods

Reservoir fluid properties were taken from an open database for the Volve field, North Sea, Norwegian sector. Machine learning methods formed the basis for determining reservoir properties of fluids and calculating technological development indicators.

## Keywords

reservoir properties, oil and gas, machine learning, artificial intelligence

## For citation

Freiman O.A., Eremin N.A. Development of a methodology for predicting reservoir properties of oil using machine learning methods. Exposition Oil Gas, 2023, issue 7, P. 118–120. (In Russ). DOI: 10.24412/2076-6785-2023-7-118-120

Received: 17.10.2023

В нефтегазовой отрасли для построения стратегии разработки месторождений очень важным фактором является прогнозирование пластовых данных и свойств нефти. Одним из важных факторов является снижение давления и дебита в процессе эксплуатации нефтяного месторождения. Данные факторы влияют на рентабельность и экономическую целесообразность разработки месторождения. Необходимо точно оценивать пластовые свойства нефти на разных этапах разведки и добычи нефти. Эти параметры имеют решающее значение для оценки течения поровой жидкости, эффективности скважины, моделирования скважин, испытания скважин и течения в трубах для производства и проектирования трубопроводов, транспортного оборудования и проектирования эксплуатации и разработки месторождений.

Точность измерений в разработке месторождений (например, расчеты материального баланса, подсчет запасов, предиктивный анализ данных) зависит от достоверности пластовых данных. В литературе влияние неопределенности пластовых данных на процесс разработки месторождения углеводородов рассматривалось в уравнениях материального баланса, оценках запасов углеводородов и выхода более летучих флюидов. В последние десятилетия было разработано большое количество эмпирических и полумпирических зависимостей, в основном из соответствующего уравнения состояния для прогнозирования пластовых свойств нефти. Большинство представленных корреляций установлены для одного региона, поэтому при их использовании для других районов будут получены ошибочные результаты [1–3]. Некоторые авторы связывают пластовые свойства сырой нефти с обычно трудноизмеримыми свойствами, такими как молярная масса, критическая температура и ацентрический фактор [4, 5].

Возникает необходимость в разработке методов, позволяющих оценить вязкость нефти в более широком диапазоне температурных интервалов, когда отсутствует физический образец для проведения лабораторных измерений или существуют ограничения по времени и требуются быстрые результаты, даже при наличии образцов флюида. Это продолжило путь для модификации и принятия уже существующих эмпирических корреляций в течение определенного периода времени. В последние десятилетия для оценки пластовых свойств нефти использовались различные модели машинного обучения [6–16]. Среди них искусственная нейронная сеть привлекла значительное внимание с конца 1990-х годов. При использовании алгоритмов машинного обучения следует иметь в виду, что теоретически, поскольку пластовые свойства нефти каждого типа будут различаться, необходимо тестировать различные модели машинного обучения, и одного алгоритма будет недостаточно.

Использование специального программного обеспечения для создания прогнозируемых моделей может занимать продолжительное время и потребовать значительной вычислительной мощности. Трудоемкие для расчетов эмпирические формулы не являются универсальным инструментом для определения и прогнозирования пластовых параметров для всех месторождений углеводородов. В данной работе представлена методика прогнозирования дебита нефти и расчета кривых падения пластового давления посредством использования алгоритмов машинного обучения.

Для оценки и точного моделирования была выполнена перекрестная проверка наборов

входных данных, а также представлено сравнение эффективности разных методов машинного обучения. Для оценки эффективности машинного обучения для каждого набора данных использовались диаграммы измеренных и прогнозируемых значений. Надежность и точность моделей в этом исследовании оценивались с использованием различных статистических показателей качества, включая коэффициент детерминации, среднюю квадратичную ошибку. Большинство методов корреляции основаны на данных разного происхождения, поэтому некоторые эмпирические зависимости не могут дать достаточной для эффективного прогнозирования точности.

Основная цель данного исследования заключается в разработке передовых методов машинного обучения для прогнозирования пластовых свойств нефти. Для разработки данной методики прогнозирования использовалась база данных с нефтяного месторождения Вольве в Северном море. Перед анализом данные были подготовлены, в них были удалены пропуски и аномальные значения. Подготовленная база пластовых данных составила 5 300 значений. Основными рассматриваемыми характеристиками стали потери давления на долоте, потери давления на бурительной колонне, давление в обсадной колонне, давление на штуцере, а также пластовые свойства нефти, такие как температура на устье и среднее давление.

Для анализа пластовых данных была использована кроссвалидация, данные были разделены на 5 частей (20/80, где 20 % — это тестовая выборка, а 80 % — обучающая). Из использованных алгоритмов наиболее хорошо себя проявили: random forest (случайный лес), KNN (метод ближайших соседей), метод градиентного бустинга. Метод случайного леса — алгоритм машинного обучения, заключающийся в использовании ансамбля деревьев решений. Дерево решений в свою очередь является отдельным алгоритмом, в ходе которого происходит классификация данных и разбитие их на классы по бинарной логике. В случае использования метода ближайших соседей для классификации объект присваивается тому классу, который является наиболее распространенным среди k соседей данного элемента, классы которых уже известны. В случае разработки методики в данной работе объекту присваивается среднее значение по k ближайшим к нему объектам, значения которых уже известны. Градиентный бустинг позволяет построить достоверную модель в виде ансамбля моделей прогнозирования, которые обычно представляют собой деревья решений.

Достоверность разработанной модели оценивалась с использованием следующих показателей:  $R^2$  (коэффициент детерминации), MSE и RMSE (средняя квадратичная ошибка), MAE (средняя абсолютная ошибка). Чем меньше значение ошибок и чем ближе  $R^2$  к единице, тем выше достоверность разработанной модели. Показатели представлены в таблице 1.

Результаты показывают, что каждая разработанная модель машинного обучения достигла приемлемой точности при применении к набору данных. В то же время из трех моделей можно выделить модель, построенную с использованием градиентного бустинга. Как показали результаты, модель эта превзошла другие модели во всех метриках. Можно выделить близкий к 1 коэффициент детерминации  $R^2 = 0,99$ . Средняя квадратичная ошибка составила 36,67 на тестовом наборе данных. На рисунке 1 приведен сравнительный график рассчитанных и исходных значений для падений давления.

## Итоги

В данном исследовании изучены методы машинного обучения с точки зрения создания модели прогнозирования показателей разработки месторождения. Результаты тестирования показали, что все предложенные модели являются достоверными, способствуют уменьшению временных затрат и требуемой вычислительной мощности.

## Выводы

Модель машинного обучения обеспечивает возможность прогнозирования пластовых свойств нефти в условиях неполной исходной информации. Разработанная методика может быть использована для начального и экономически выгодного прогнозирования добычи нефти в начальный период разработки, что поможет сэкономить выделенные средства и лучше оценить возможные риски.

## Литература

1. Брусиловский А.И., Ющенко Т.С. Научно обоснованный инженерный метод определения компонентного состава и PVT-свойств пластовых углеводородных смесей при неполной исходной информации // ПРОНефть. 2016. № 1. С. 68–74.
2. Дунюшкин И.И., Мищенко И.Т., Елисеева Е.И. Расчеты физико-химических свойств пластовой и промысловой нефти и воды. М.: Нефть и газ, 2004. 448 с.
3. Ющенко Т.С., Брусиловский А.И. Эффективный метод построения и адаптации PVT-моделей пластовых флюидов газоконденсатных месторождений и газовых шапок нефтегазоконденсатных залежей // Нефтяное хозяйство. 2015. № 1. С. 56–60.
4. Al-Marhoun M.A., Duffuaa S.O., Abu-Khamsin S.A. Viscosity correlations for Saudi Arabian crude oils. Middle East oil show, March 1987, Bahrain, SPE-15720-MS. (In Eng).
5. Al-Fattah S.M. Artificial-Intelligence technology predicts relative permeability of giant carbonate reservoirs. SPE Offshore Europe Conference, 4–7 September 2007, Aberdeen, Scotland, U.K, SPE 109018. (In Eng).

Табл. 1. Показатели достоверности модели  
Tab. 1. Model reliability indicators

	MSE	RMSE	MAE
RF	1 510,877	38,87	15,65
KNN	2 697,764	51,94	30,46
GB	1 344,689	36,67	14,16

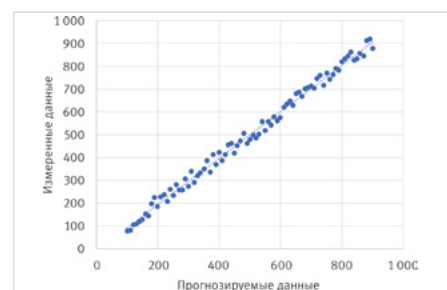


Рис. 1. Визуализация точности модели GB (градиентный бустинг)

Fig. 1. Visualization of GB (gradient boosting) model accuracy

6. Naseri A., Yousefi S.H., Sanaei A., Ghahsheikhlou A.A. A neural network model and an updated correlation for estimation of dead crude oil viscosity. *Brazilian Journal of Petroleum and Gas*, 2012, Vol. 6, issue 1, P. 31–41. (In Eng).
7. Hadavimoghaddam F., Mishchenko I. Introducing A new correlation method to predict PVT properties of Russian and Iranian crude oils. Student Technical Conference (STC), 7–9 November 2018, Freiberg. (In Eng).
8. Jain D., Bihani A.D. Crude oil viscosity correlations: a novel approach for Upper Assam basin. *Petrotech 2014*, New Delhi, India. (In Eng).
9. Ghorbani B., Ziabasharhagh M., Amidpour M. A hybrid artificial neural network and genetic algorithm for predicting viscosity of Iranian crude oils. *Journal of natural gas science and engineering*, 2014, Vol. 18, P. 312–323. (In Eng).
10. Ghorbani H., Wood D.A., Choubineh A., Mohamadian N., Tatar A., Farhangian H., Nikooy A. Performance comparison of bubble point pressure from oil PVT data: several neurocomputing techniques compared. *Experimental and computational multiphase flow*, 2020, Vol. 2, issue 2, P. 225–246. (In Eng).
11. Фрейман О.А. Особенности освоения ресурсов нефти и газа в отложениях триаса // *Научный взгляд в будущее*. 2021. Том 1. № 23. С. 53–57.
12. Фрейман О.А. Проблемы и перспективы освоения арктической нефти // *Природопользование и устойчивое развитие регионов России*. Пенза: ПГАУ, 2022. С. 87–92.
13. Фрейман О.А. Особенности освоения нефти на арктическом шельфе // *Передовое развитие современной науки: опыт, проблемы, прогнозы*. Петрозаводск: Новая наука, 2022. С. 172–177.
14. Фрейман О.А. Актуальность применения технологии секвестрации углекислого газа. *Природопользование и устойчивое развитие регионов России*. Пенза: ПГАУ, 2023. С. 95–97.
15. Фрейман О.А. Анализ применения статистических и компьютерных методов в нефтегазовой промышленности в условиях неполной исходной информации // *Современные исследования: теория и практика*. Петрозаводск: Новая наука, 2022. С. 334–337.
16. Фрейман О.А. Математическое моделирование и машинное обучение в нефтегазовой промышленности // *Формирование профессиональной направленности личности специалистов — путь к инновационному развитию России*. Пенза: ПГАУ, 2022. С. 226–229.

## ENGLISH

### Results

In this study, machine learning methods have been investigated in terms of creating a model for predicting field development performance. The results of testing showed that all proposed models are reliable, contribute to the reduction of time costs and required computing power.

### References

1. Brusilovskiy A.I., Yushchenko T.S. Two-phase deposits: methodology approach to the identification of composition and pvt properties of reservoir hydrocarbon fluids using limited initial information. *ProNeft*, 2016, issue 1, P. 68–74. (In Russ).
2. Dunyushkin I.I., Mishchenko I.T., Eliseeva E.I. Calculations of physicochemical properties of reservoir and field oil and water. *Moscow: Oil and Gas*, 2004, 448 p. (In Russ).
3. Yushchenko T.S., Brusilovsky A.I. Effective method for creating and adapting PVT-model of reservoir fluid of gas-condensate deposits and gas cap for oil-gas-condensate reservoirs. *Oil industry*, 2015, issue 1, P. 56–60. (In Russ).
4. Al-Marhoun M.A., Duffuaa S.O., Abu-Khamsin S.A. Viscosity correlations for Saudi Arabian crude oils. *Middle East oil show, March 1987, Bahrain, SPE-15720-MS*. (In Eng).
5. Al-Fattah S.M. Artificial-Intelligence technology predicts relative permeability of giant carbonate reservoirs. *SPE Offshore Europe Conference*, 4–7 September 2007, Aberdeen, Scotland, U.K., SPE 109018. (In Eng).
6. Naseri A., Yousefi S.H., Sanaei A., Ghahsheikhlou A.A. A neural network model and an updated correlation for estimation of dead crude oil viscosity. *Brazilian Journal of Petroleum and Gas*, 2012, Vol. 6, issue 1, P. 31–41. (In Eng).
7. Hadavimoghaddam F., Mishchenko I. Introducing A new correlation method to predict PVT properties of Russian and Iranian crude oils. Student Technical Conference (STC), 7–9 November 2018, Freiberg. (In Eng).
8. Jain D., Bihani A.D. Crude oil viscosity correlations: a novel approach for Upper Assam basin. *Petrotech 2014*, New Delhi, India. (In Eng).
9. Ghorbani B., Ziabasharhagh M., Amidpour M. A hybrid artificial neural network and genetic algorithm for predicting viscosity of Iranian crude oils. *Journal of natural gas science and engineering*, 2014, Vol. 18, P. 312–323. (In Eng).
10. Ghorbani H., Wood D.A., Choubineh A., Mohamadian N., Tatar A., Farhangian H., Nikooy A. Performance comparison of bubble point pressure from oil PVT data: several neurocomputing techniques compared. *Experimental and computational multiphase flow*, 2020, Vol. 2, issue 2, P. 225–246. (In Eng).
11. Freiman O.A. Features of the development of oil and gas resources in the triassic sediments. *Scientific look into the future*, 2021, Vol. 1, issue 23, P. 53–57. (In Russ).
12. Freiman O.A. Problems and prospects of Arctic oil development. *Environmental management and sustainable development of Russian regions*. Penza: Penza state agrarian university, 2022, P. 87–92. (In Russ).

### Conclusions

The machine learning model provides the possibility of predicting oil reservoir properties in conditions of incomplete initial information. The developed methodology can be used for initial and cost-effective forecasting of oil production in the initial period of development, which will help to save allocated funds and better assess possible risks.

13. Freiman O.A. Main features of oil development on the Arctic shelf. *Advanced development of modern science: experience, problems, forecasts*. Petrozavodsk: Novaya nauka, 2022, P. 172–177. (In Russ).
14. Freiman O.A. Relevance of application of carbon dioxide sequestration technology. *Nature management and sustainable development of regions of Russia*. Penza: Penza state agrarian university, 2023, P. 95–97. (In Russ).
15. Freiman O.A. Analysis of the application of statistical and computer methods in the oil and gas industry in conditions of incomplete initial information. *Petrozavodsk: Novaya nauka*, 2022, P. 334–337. (In Russ).
16. Freiman O.A. Mathematical modeling and machine learning in the oil and gas industry. *Formation of professional orientation of specialists' personalities – the path to innovative development of Russia*. Penza: Penza state agrarian university, 2022, P. 226–229. (In Russ).

## ИНФОРМАЦИЯ ОБ АВТОРАХ | INFORMATION ABOUT THE AUTHORS

**Фрейман Олег Александрович**, аспирант, САФУ имени М.В. Ломоносова, Архангельск, Россия  
**Для контактов:** [frejman.o@edu.narfu.ru](mailto:frejman.o@edu.narfu.ru)

**Еремин Николай Александрович**, д.т.н., заведующий Аналитическим центром, Институт проблем нефти и газа РАН, Москва, Россия; профессор, Губкинский университет, Москва, Россия; научный руководитель, САФУ имени М.В. Ломоносова, Архангельск, Россия

**Freiman Oleg Alexandrovich**, postgraduate, Northern (Arctic) Federal University, Arkhangelsk, Russia  
**Corresponding author:** [frejman.o@edu.narfu.ru](mailto:frejman.o@edu.narfu.ru)

**Eremin Nikolai Alexandrovich**, doctor of engineering sciences, head of the analytical center, Oil and Gas Research Institute RAS, Moscow, Russia; professor, Gubkin University, Moscow, Russia; scientific director of the department of drilling of wells and development of oil and gas fields of the Northern Arctic Federal University named after M.V. Lomonosov, Arkhangelsk, Russia