

А. П. Саматова^{1,2✉}, И. В. Филимонова^{1,2}

Методические основы прогнозирования цен на нефть

¹Новосибирский государственный университет,
г. Новосибирск, Российская Федерация

²Институт нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН,
г. Новосибирск, Российская Федерация
e-mail: a.samatova@g.nsu.ru

Аннотация. Прогнозы цен на нефть оказывают значимое влияние на экономику большинства стран. Они используются в стратегических документах, обеспечивают уверенность инвесторов, позволяют производителям планировать свою деятельность. В связи с этим чрезвычайно важной задачей является выявление способа, обеспечивающего высокую точность прогноза. Целью исследования является обзор методов прогнозирования цен на нефть. Показано, что существующие методы прогнозирования цен на нефть можно разделить на количественные и качественные методы. Выявлены основные ограничения существующих подходов, показаны преимущества и недостатки наиболее популярных методов. Представлена классификация методов машинного обучения. Показано, что количественные и качественные методы можно применять вместе с целью обогащения исследования и учета влияния факторов, включение которых в модели ранее было затруднено.

Ключевые слова: прогнозирование, цены на нефть, эконометрические модели, машинное обучение

A. P. Samatova^{1,2✉}, I. V. Filimonova^{1,2}

Methodological foundations of oil price forecasting

¹Novosibirsk State University, Novosibirsk, Russian Federation

²Trofimuk Institute of Petroleum Geology and Geophysics SB RAS,
Novosibirsk, the Russian Federation
e-mail: a.samatova@g.nsu.ru

Abstract. Oil price forecasts have a significant impact on the economies of most countries. They are used in strategic documents, provide investors with confidence, and allow manufacturers to plan their activities. In this regard, it is extremely important to identify a method that ensures high forecast accuracy. The purpose of the study is to review the methods of forecasting oil prices. It is shown that the existing methods of forecasting oil prices can be divided into quantitative and qualitative methods. The main limitations of the existing approaches are revealed, the advantages and disadvantages of the most popular methods are shown. The classification of machine learning methods is presented. It is shown that quantitative and qualitative methods can be applied together in order to enrich the research and take into account the influence of factors that were previously difficult to include in the models.

Keywords: forecasting, oil prices, econometric models, machine learning

Введение

Все страны в той или иной мере являются участниками нефтяного рынка. От колебаний цен на нефть зависят доходы и расходы государств, предприятий,

использующих нефть в качестве топлива и сырья, инвесторов и других участников. На данный момент в России реализуется и планируется к реализации множество проектов, направленных на развитие нефтегазового комплекса, в том числе в Восточной Сибири и на Дальнем Востоке [1-3], эффект от реализации которых зависит в том числе от цен на нефть. При этом динамика цен на нефть характеризуется высокой волатильностью, цены способны измениться достаточно сильно за относительно короткий период времени. Так, с июня 2022 г. по июнь 2023 г. цены на нефть марки Brent упали с 117,69 до 74,98 долларов за баррель. То есть, всего за год цены на нефть упали на 36,29%, что значительно снизило доходы стран-экспортеров нефти. Динамика цен на нефть зависит от множества факторов, причем наиболее ярко на цены влияют геополитические события, часто служащие причинами резкого роста или падения цен (рис. 1). Таким образом, возникает необходимость хеджирования рисков, связанных с изменениями цен на нефть. Неотъемлемой частью этого процесса является прогнозирование цен на нефть.



Рис. 1. Динамика цен на нефть марки WTI, долл. за барр. и основные события, объясняющие снижения и повышения цен

Источник: составлено автором по данным Federal Reserve Bank of St. Louis (FRED)

Подходы, применяемые при прогнозировании цен на нефть, весьма разнообразны и используются для решения разных задач. Так, например, [4] использовали модификацию модели ARIMA для прогнозирования цен биржевых товаров, в том числе нефти. [5] использовали модификацию модели GARCH, чтобы оценить влияние геополитических рисков на волатильность цен нефти. А в [6] используется модель коррекции ошибок для моделирования долгосрочных равновесных отношений между ценами на нефть и природный газ. Таким образом,

понимание особенностей, преимуществ и недостатков тех или иных подходов позволит более точно определить модель при выполнении определенной задачи и добиться лучших результатов.

Целью данного исследования является обзор методов прогнозирования цен на нефть для упрощения ориентирования в системе существующих подходов.

Для достижения цели был поставлен ряд задач:

4. Изучить теоретические основы прогнозирования цен на нефть;
5. Систематизировать существующие методы прогнозирования цен на нефть;
6. Выявить особенности, преимущества и недостатки крупных групп методов;
7. Сформировать выводы по полученным результатам.

Результаты и обсуждение

Выделяют *количественные* и *качественные* методы прогнозирования цен на нефть. Количественные методы основаны на использовании математических моделей в процессе прогнозирования, обычно в результате получается конкретное значение будущей цены или других связанных прогнозируемых параметров (волатильности, темпов роста и т.д.). Количественные методы намного более распространены в исследованиях, среди них выделяются эконометрический анализ, машинное обучение и смешанные подходы, призванные соединить преимущества первых двух групп методов.

Эконометрические модели совсем недавно были доминирующим методом прогнозирования и до сих пор активно применяются в исследованиях. К их преимуществам относится высокая интерпретируемость, возможность получить конкретные оценки влияния факторов (коэффициенты модели). Такие модели прекрасно справляются с выявлением линейных взаимосвязей, многие из них относительно просты в реализации. Одним из наиболее распространенных ответвлений эконометрического анализа является анализ временных рядов, в частности модели ARIMA, GARCH, VAR и их модификации. Однако анализ с помощью эконометрических моделей ограничен необходимостью выполнения строгих предположений, например, о стационарности рядов, экзогенности предикторов, нормальности ошибок и т.д. Помимо этого, большинство моделей очень чувствительны к наличию пропусков и выбросов в данных, что требует от исследователя тщательной работы с данными перед построением модели. Одним из наиболее существенных ограничений является предположение о линейности взаимосвязей.

Стремление зайти дальше и выявить нелинейные взаимосвязи в данных привело к развитию и росту популярности машинного обучения. Нелинейные модели позволяют избавиться от основных ограничений эконометрических методов, они хорошо справляются с наличием выбросов и пропусков. Однако и они имеют свои ограничения. Так, например, машинное обучение лучше всего работает при использовании большого количества данных, получение которых затруднено при анализе энергетических рынков. В случае, если исследователь хо-

чет учесть влияние фундаментальных факторов, таких как уровень добычи или запасов, ему придется использовать данные с ежемесячной периодичностью, так как данные публикуются не так часто. Таким образом, обучение будет проводиться на очень ограниченной, по меркам машинного обучения, выборке. В результате может возникнуть проблема переобучения, при возникновении которой модель излишне подстраивается под обучающую выборку и плохо обобщает закономерности на тестовой выборке. Справиться с этим помогает регуляризация, использование неглубоких деревьев и ранняя остановка обучения.

На рисунке 2 представлена распространенная классификация методов машинного обучения. Обучение с учителем предполагает использование размеченных данных, содержащих значение целевой переменной, которое модель будет обучаться предсказывать. Обучение без учителя применяется в случае, если «правильный ответ» заранее неизвестен и требуется найти взаимосвязи в данных. Часто обучение без учителя применяется для решения задачи кластеризации. Еще одним подходом к обучению моделей является обучение с подкреплением, в ходе которого модель взаимодействует с окружающей средой и получает вознаграждения в случае выполнения правильных действий, стремясь максимизировать общий выигрыш. Этот способ обучения особенно популярен в сферах, где присутствует высокая степень вариативности возможных правильных и неправильных действий, например, в робототехнике или транспортных системах.

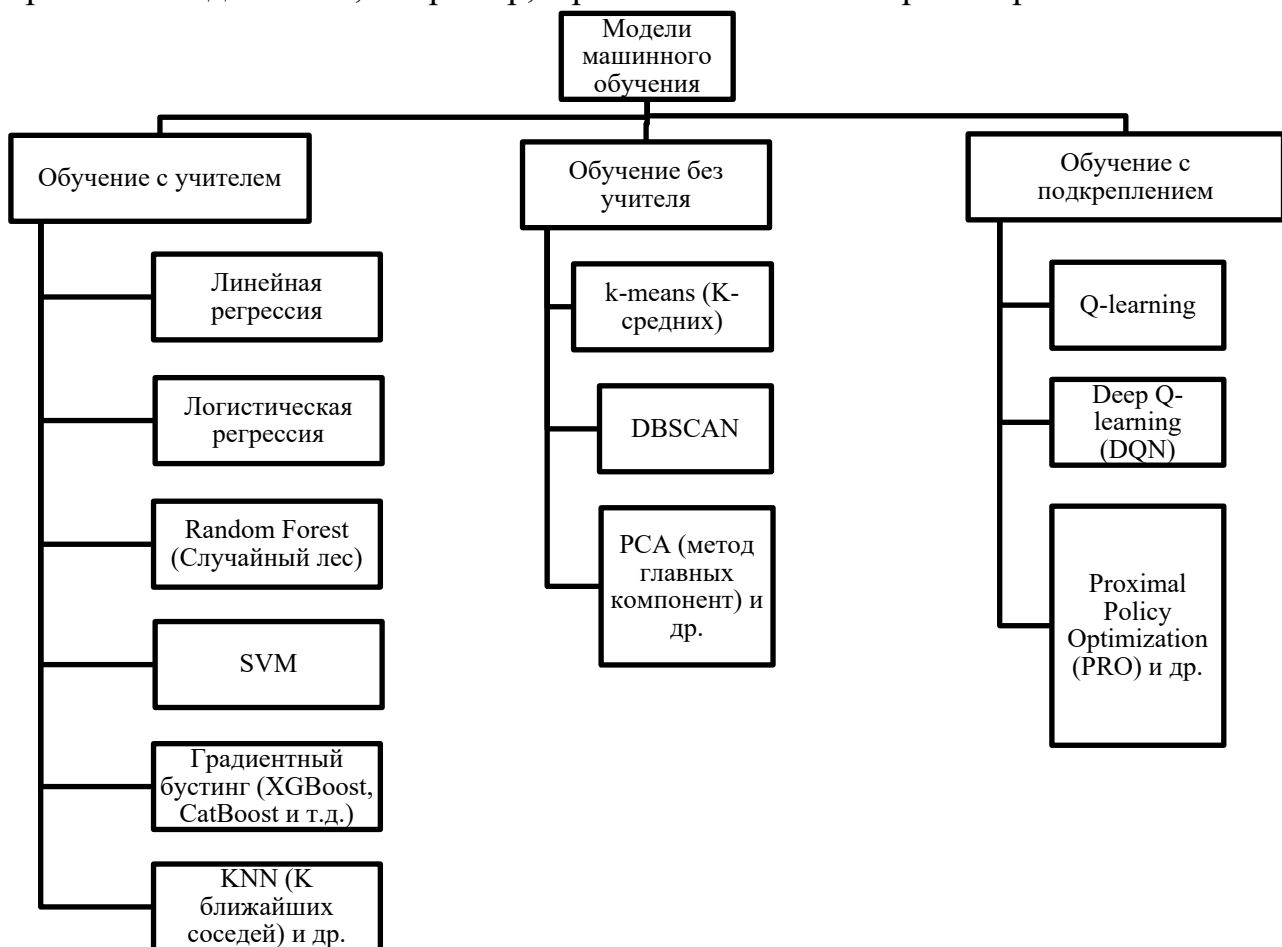


Рис. 2. Классификация моделей машинного обучения

Источник: составлено автором

Применительно к задаче прогнозирования цен на нефть решается задача регрессии, которая относится к методам обучения с учителем. В последние годы большинство исследований показывает, что эффективнее всего задачу прогнозирования цен на нефть решают алгоритмы глубокого обучения, в частности, модель LSTM (Long Short-Term Memory). Однако, несмотря на высокую эффективность таких алгоритмов, существует ряд ограничений их использования. Во-первых, глубокое обучение требует еще большего объема данных для эффективной работы, чем классическое машинное обучение. В контексте прогнозирования цен на нефть с помощью LSTM проблема решается путем использования только высокочастотных данных (например, данные о прошлых ценах на нефть, ценах на золото, серебро и т.д.). Однако в случае, если исследователь хочет оценить влияние и других факторов, данные о которых публикуются реже, ему лучше обратиться к классическому машинному обучению. Другим ограничением является низкая интерпретируемость таких моделей, что может быть критичным ограничением при исследовании влияния факторов на прогноз. Таким образом, глубокое обучение имеет свои преимущества и недостатки и часто применяется в исследованиях, ставящих целью получение прогноза, а не интерпретацию влияния факторов.

Наиболее высокой эффективностью среди классических методов машинного обучения характеризуются модели, основанные на градиентном бустинге, такие как XGBoost и CatBoost. Эти модели основаны на использовании большого количества небольших «слабых» моделей (обычно в качестве таких моделей используются деревья решений), обучающихся последовательно и стремящихся исправить ошибки предыдущих моделей. Таким образом, ошибки отдельных деревьев меньше влияют на результат и точность прогноза повышается.

Смешанные подходы встречаются реже, однако используются для устранения ограничений эконометрических методов и методов машинного обучения. Это позволяет совместить высокую точность прогноза моделей машинного обучения и интерпретируемость эконометрических моделей.

Качественные подходы применяются для учета факторов, влияния которых сложно моделировать: влияние войн, новостей, доступности рынка и т.д. Распространенным качественным методом, часто применяемым в контексте нефтяных цен, является метод Дельфи, основанный опросах экспертов. Другим методом, основанным на опросах экспертов, является метод нечеткого ранжирования предпочтений PROMETHEE, примененный в работе [7]. Результаты качественных исследований часто используются в количественных исследованиях: так, в результате анализа текстов новостей формируются индексы геополитической напряженности, неопределенности экономической политики, климатической неопределенности и т.д., которые могут использоваться при моделировании.

Заключение

В ходе исследования были проанализированы методические подходы к прогнозированию цен на нефть. Было выявлено, что количественные подходы более

распространены в литературе, посвященной прогнозированию цен на нефть. Количественные методы можно разделить на эконометрические модели, модели машинного обучения и смешанные подходы. Машинное обучение набирает популярность за счет способности выявлять нелинейные взаимосвязи и большей гибкости моделей, однако требуют большого количества данных и уступают эконометрическим моделям в интерпретируемости. Смешанные подходы позволяют объединить преимущества обеих групп методов, однако это направление пока менее развито. При этом качественные методы позволяют учитывать сложномоделируемые факторы, что существенно обогащает анализ, однако ограничением данного метода является субъективность источников. Таким образом, можно сделать вывод о необходимости тщательного подбора наиболее подходящей модели в контексте решаемой проблемы, при этом внимание стоит обращать не только на моделируемый параметр, но и на то, с какой периодичностью можно получить данные, каким окажется итоговый размер выборки, важна ли интерпретируемость прогноза. Результаты исследования могут быть полезны при определении наиболее эффективного механизма прогнозирования цен на нефть.

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 23-78-10157, <https://rscf.ru/project/23-78-10157/>.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Eder L.V., Filimonova I.V., Provornaya I.V., Komarova A.V., Nikitenko S.M. New directions for sustainable development of oil and gas industry of Russia: innovative strategies, regional smart specializations, public-private partnership // В сборнике: 17TH INTERNATIONAL MULTIDISCIPLINARY SCIENTIFIC GEOCONFERENCE SGEM 2017. Conference proceedings. – 2017. – С. 365-372.
2. Эдер Л.В., Саблин К.С., Проворная И.В. Научные подходы к обоснованию приоритетных инновационно-технологических направлений пространственной специализации ресурсных регионов России // Фундаментальные исследования. – 2017. – № 5. – С. 220-224.
3. Эдер Л.В., Проворная И.В. Основные направления инновационного развития нефтегазовой промышленности России // В сборнике: Инновационный потенциал экономики России: состояние и перспективы. Алексеев А.В., Казанцева Л.К. Сборник научных трудов. Институт экономики и организации промышленного производства СО РАН. – Новосибирск, 2013. – С. 165-183.
4. Gupta H. et al. Forecasting Commodity Prices using Machine Learning //International Journal of Scientific Research in Science and Technology. – 2024. – Vol. 11. – №1. – DOI: 10.32628/IJSRST52411110.
5. Wang L. et al. Forecasting crude oil volatility with geopolitical risk: Do time-varying switching probabilities play a role? //International Review of Financial Analysis. – 2021. – Vol. 76. – №101756. – DOI: 10.1016/j.irfa.2021.101756.
6. Brigida M. The switching relationship between natural gas and crude oil prices //Energy Economics. – 2014. – Vol. 43. – P. 48-55. – DOI: 10.1016/j.eneco.2014.01.014.
7. Abdullah L. et al. Preorder of Factors Affecting Oil Prices: Fuzzy PROMETHEE Approach //Journal of Mathematics. – 2022. – Vol. 2022. – №1. – P. 2766945.

© А. П. Саматова, И. В. Филимонова, 2025