

УДК 004

## Исследование методов прогнозирования добычи нефти

А. А. Сорокин, А. В. Гурко

Санкт-Петербургский Горный Университет императрицы Екатерины II

E-mail: s211920@stud.spmi.ru

### Аннотация

*В статье анализируются методы прогнозирования добычи нефти с применением нейронных сетей, выполняется компьютерный эксперимент оценки вероятности успешной добычи нефти на основе анализа геофизических параметров. Представлена архитектура нейронной модели, приведены результаты её обучения и тестирования. Направлением дальнейших исследований является расширение перечня входных параметров и обучение модели на других видах алгоритмов машинного обучения для повышения прогностических способностей модели*

### Введение

В нефтедобыче традиционно применяются методы гидродинамического моделирования, статистического анализа и эмпирические зависимости для прогнозирования дебита скважин и динамики добычи. Однако эти подходы часто ограничены в точности и адаптивности, особенно при учёте сложных геологических условий и временных изменений параметров разработки. В условиях динамично развивающейся нефтегазовой отрасли одной из ключевых задач является обеспечение стабильности и эффективности процессов добычи нефти.

Современные методы машинного обучения предоставляют новые возможности для повышения точности и оперативности прогнозирования. С развитием вычислительных технологий и машинного обучения методы прогнозирования добычи нефти претерпели значительные изменения. Одним из таких методов является использование нейросетевых моделей, которые учитывают сложные взаимосвязи между параметрами работы месторождения. Перспективными являются рекуррентные нейронные сети, такие как LSTM (long short-term memory)<sup>1</sup>, которые эффективно моделируют временные зависимости. Замена традиционных подходов методами машинного обучения позволяют значительно повысить точность и оперативность прогнозирования. Например, в [1] исследуется использование искусственного интеллекта для автоматизации процессов разведки и добычи нефти.

В данном исследовании рассматриваются методы, такие как случайный лес и нейронные сети, для повышения точности прогноза и оптимизации процессов бурения. Вершинин В. Е.

и Пономарев Р. Ю. показали, что нейросетевые технологии могут быть использованы для прогнозирования работы скважин в условиях нестационарного заводнения, что особенно актуально для оптимизации добычи в реальном времени. Это исследование подтверждает, что нейросетевые модели позволяют значительно сократить время вычислений, что критически важно для оперативного управления процессами [2]. В [3] показано, что нейросетевые модели подходят для анализа временных рядов, что позволяет более точно прогнозировать параметры работы добывающих скважин.

Другим подходом является использование компьютерного моделирования для прогнозирования добычи, например в патенте № 2794707 [4] предлагается метод, который сочетает геофизические, гидродинамические и тепловые характеристики для более точных прогнозов объема добычи углеводородов. В [5, 6] подчеркивается важность применения машинного обучения для прогнозирования добычи нефти с использованием исторических данных. Использование описанных технологий позволяет не только ускорить процесс анализа, но и более точно предсказать уровни добычи, что является важным для принятия оперативных решений в процессе разработки месторождения. В статье [7] с помощью нейросетевого подхода улучшается точность прогнозов по добыче. В [8] показано, что применение машинного обучения позволяет повысить эффективность бурения, за счет прогнозирования механических нагрузок на оборудование. В работе [9] предложены методики построения математической модели, основанной на экспертных, что позволяет обобщить данные и использовать их для дальнейшего анализа эффективности процессов. Исследователи Негаш

<sup>1</sup> LSTM layer. [https://keras.io/api/layers/recurrent\\_layers/lstm/](https://keras.io/api/layers/recurrent_layers/lstm/)  
(Дата обращения: 13.05.2025)

Б. М. и Ява А. Д. в [10] предлагают использовать нейронные сети для прогнозирования добычи в условиях заводнения скважин. В [11] рассматриваются проблемы и возможности искусственного интеллекта в горнодобывающем секторе промышленности.

### **Применение интеллектуальной системы прогнозирования добычи нефти**

Для выполнения прогнозирования добычи нефти по историческим геофизическим данным представлен новый метод, реализованный в виде интеллектуальной системы, использующей полносвязную нейронную сеть. Выбор архитектуры многослойной нейронной сети прямого распространения для построения интеллектуальной системы прогнозирования добычи нефти обусловлен её способностью моделировать сложные нелинейные зависимости между большим числом входных геофизических параметров и целевым показателем.

В настоящий момент общепринятой схемой обучения нейронной сети является применение нелинейной функций активации ReLU (rectified Linear Unit)<sup>2</sup> во внутренних слоях или сигмоидальной функции<sup>3</sup> на выходе модели. Математически функция активации ReLU определяется следующим образом:

$$ReLU(x) = \max(0, x), \quad (1)$$

где  $\max$  – функция, возвращающая максимальное значение из двух.

Математически сигмовидная функция активации определяется следующим образом:

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad (2)$$

где  $x$  – входное значение нейрона (суммарное взвешенное значение после линейной комбинации входов),  $\sigma(x)$  – выходное значение нейрона, интерпретируемое как вероятность,  $e$  – основание натурального логарифма.

Функция используется для преобразования выходного значения нейрона в вероятность, т.е. вероятность того, что входное значение относится к классу 1, если мы работаем с задачей бинарной классификации. Если значение сигмовидной функции близко к 1, то вероятность того, что входное значение относится к классу 1, высока. Если значение близко к 0, то вероятность того, что входное значение относится к классу 1, низкая.

После вычисления значения функции потерь начинается этап распространения ошибки

от выходного слоя к входному слою. Алгоритм реализует метод градиентного спуска, позволяя пошагово корректировать значения весов и смещений в слоях сети на основе вычисленной ошибки между фактическим и предсказанным значениями. Обратное распространение включает последовательное применение правила цепного дифференцирования от выходного слоя к входным.

Процесс обновления весов и смещений в нейронной сети представляет собой ключевую часть алгоритма обучения. После того, как вычисляются производные функций потерь, параметры модифицируются в направлении, которое уменьшает значение функции потерь. Этот процесс происходит при помощи оптимизатора Adam (adaptive moment estimation)<sup>4</sup>. Он реализует адаптивное изменение шага обучения для каждого параметра на основе оценки первого и второго моментов градиентов. Алгоритм Adam корректирует скорость обучения для каждого параметра индивидуально на основе первого момента (математического ожидания) и второго момента (дисперсии) градиента. Это позволяет избежать резких колебаний при обучении и ускорить сходимость.

Для решения задачи программирования выбрана нейронная сеть FNN, состоящая из пяти слоев. 128 признаков, выходной признак 1. Также использованы функции активации - сигмоидная и ReLU. Фрагмент программного кода представлен ниже:

```
model = Sequential([
    Dense(128, activation='relu',
input_shape=(X_train.shape[1],)),
    BatchNormalization(),
    Dropout(0.3),
```

Данный слой использует функцию активации согласно формуле (1).

```
Dense(256, activation='relu',
BatchNormalization(),
Dropout(0.4),
Dense(128, activation='relu',
BatchNormalization(),
Dropout(0.3),
Dense(64, activation='relu',
BatchNormalization(),
```

Данный слой использует функцию активации согласно формуле (2).

```
Dense(1, activation='sigmoid') ])
```

```
Использован оптимизатор Adam.
optimizer = Adam(learning_rate=0.001) //
model.compile(optimizer=optimizer, loss='mse',
metrics=['mae',
tf.keras.metrics.RootMeanSquaredError()])
early_stopping = EarlyStopping(monitor='val_loss',
patience=20,
restore_best_weights=True)
```

<sup>2</sup> Relu layer. [https://keras.io/api/layers/activation\\_layers/relu/](https://keras.io/api/layers/activation_layers/relu/) (Дата обращения: 13.05.2025)

<sup>3</sup> Sigmoid function. <https://keras.io/api/layers/activations/> (Дата обращения: 13.05.2025)

<sup>4</sup> Adam. <https://keras.io/api/optimizers/adam/> (Дата обращения: 13.05.2025)

### Предобработка и анализ данных

Для обучения и тестирования интеллектуальной системы прогнозирования использовался табличный набор данных, предоставленный нефтедобывающей компанией. Представленный набор данных состоит из набора следующих геофизических характеристик: географическая широта и долгота точки исследования, глубина скважины, пористость породы, пластовое давление. Для выполнения

прогнозирования по заданному набору характеристик, была добавлена целевая переменная SuccessProb, которая описывает вероятность добычи нефти в рассматриваемой точке. На этапе предобработки реализована визуализация распределений, чтобы предварительно проанализировать потенциальные зависимости между признаками. Визуализация исходных данных представлена на рисунке 1.

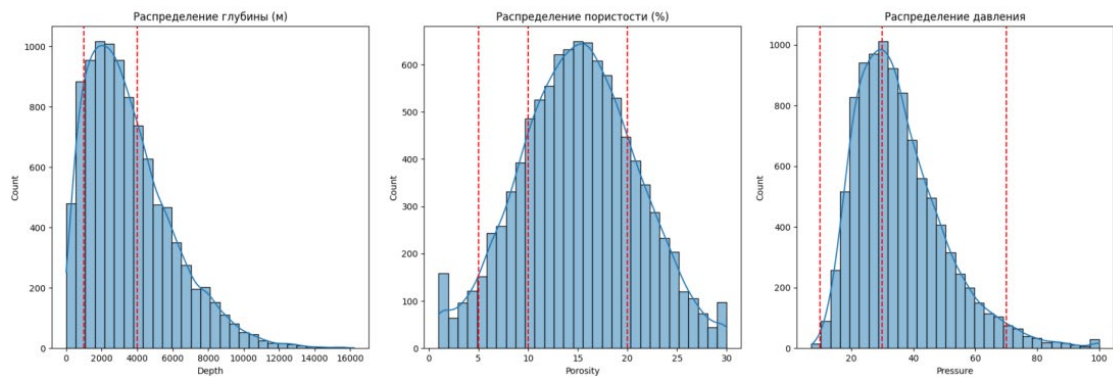


Рисунок 1 – Визуализация исходных данных [материал автора]

Для каждого из параметров были выделены пороговые значения, отражающие геологическую и производственную значимость интервалов.

По пористости выделены границы 5%, 10% и 20%, соответствующие общепринятой классификации пород по степени продуктивности. Это позволяет визуально оценить долю перспективных и неперспективных коллекторов в выборке.

Аналогично, на гистограмме давления пороговые значения в 10, 30 и 70 условных единиц обозначают зоны низкого, среднего и высокого пластового давления, что даёт возможность оценить условия разработки и необходимость использования дополнительных технологий.

На графике распределения глубин с помощью вертикальных линий акцентирован интервал 1000–4000 м., где сосредоточена основная масса данных, что позволяет выделить рабочую глубину основной выборки.

Такое представление данных способствует быстрой интерпретации и выявлению возможных аномалий или закономерностей. Чтобы данные корректно обрабатывались, их необходимо привести к виду  $[0,1]$ , чтобы они были в едином масштабе. Каждый признак был линейно преобразован по формуле:

$$x_{\text{норм}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (3)$$

где  $x$  – исходное значение признака,  $x_{\min}$  и  $x_{\max}$  – минимальное и максимальные значения признака,  $x_{\text{норм}}$  – нормализованное значение, лежащее в диапазоне от 0 до 1.

Визуализация нормализации данных на основе формулы (3) представлена на рисунке 2.

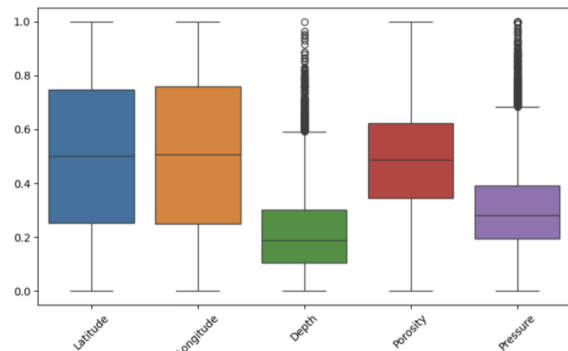


Рисунок 2 – Нормализация исходных данных [материалы авторов]

В процессе создания, обучения и тестирования модели нейронной сети было произведено разделение данных на обучающую (80%) и тестовую (20%) выборки.

После нормализации и разбиения данных на выборки они становятся готовыми для обучения модели нейронной сети.

### Визуализация результатов обучения модели

Оценка точности модели производилась на тестовой выборке. На рисунке 3 представлены три ключевые визуализации, отражающие качество работы обученной нейронной сети. Первый график иллюстрирует соответствие между истинными и предсказанными значениями целевой переменной – чем ближе точки к диагонали  $y = x$ , тем выше точность модели.

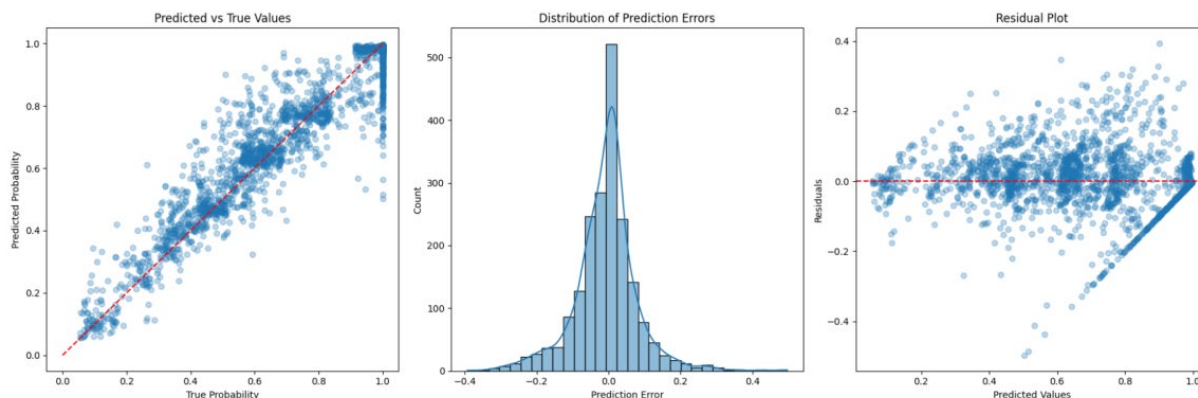


Рисунок 3 – Качество работы обученной нейронной сети [материалы авторов]

Второй график – гистограмма ошибок предсказания – демонстрирует распределение отклонений и позволяет оценить наличие систематических ошибок; её симметричная форма с пиком вблизи нуля указывает на стабильную работу модели. Третий график – график остатков – показывает, насколько равномерно распределены ошибки по диапазону предсказанных значений. Отсутствие ярко выраженного тренда и концентрация точек вдоль нулевой линии говорят о хорошем качестве предсказания без выраженного смещения.

Такая компоновка позволяет в сжатом виде оценить точность модели, характер ошибок и потенциальные проблемы, связанные с пере- или недообучением.

Рисунок 4 демонстрирует взаимосвязь между основными геологическими параметрами (глубина, давление, пористость) и вероятностью успешного бурения, предсказанной нейронной сетью. Все визуализации построены по единому принципу: положение точек на графиках определяется парой признаков, а цвет — предсказанным значением вероятности успеха.

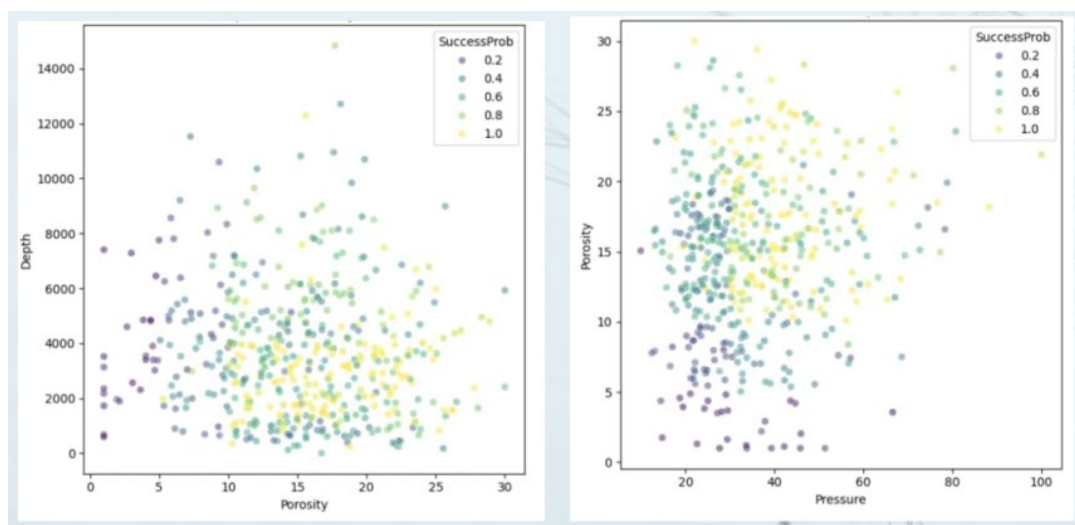


Рисунок 4 – Взаимосвязь геологических параметров [материалы авторов]

Использование цветовой палитры viridis позволяет визуально оценить уровень прогнозируемого успеха: от низкого (фиолетовые и тёмные оттенки) к высокому (жёлто-зелёные оттенки). Такая форма представления делает возможным быстрое выявление благоприятных зон сочетания параметров. Каждая точка соответствует конкретной скважине из выборки (500 случайных наблюдений), что обеспечивает баланс между плотностью информации и наглядностью графика.

### Оценка качества модели

Для оценки качества модели используется среднеквадратичная ошибка (MSE). Большее значение среднеквадратического отклонения показывает больший разброс значений в представленном множестве со средней величиной множества; меньшее значение, соответственно, показывает, что значения в множестве сгруппированы вокруг среднего значения. Такой вид оценки качества удобно использовать для

выявления аномалий при обучении и работе модели нейронной сети.

Математически среднеквадратичная ошибка (MSE) определяется следующим образом:

$$MSE = \frac{1}{m} \sum_{i=1}^m (\gamma_i - \gamma_{ist_i})^2, \quad (4)$$

где  $m$  – число объектов в обучающей выборке,  $\gamma_i$  – истинное значение,  $\gamma_{ist}$  – предсказанное значение моделью.

Дополнительно, для контроля качества работы модели используется средняя абсолютная ошибка MAE и корень из среднеквадратической ошибки RMSE. Средняя абсолютная ошибка рассчитывается как среднее абсолютных разностей между целевыми значением и

значением, предсказанным моделью на данном обучающем примере в процессе обучения:

$$MAE = \frac{1}{N} \sum_{i=1}^m |\gamma_i - \gamma_{ist_i}|, \quad (5)$$

где  $m$  – число объектов в обучающей выборке,  $\gamma_i$  – истинное значение,  $\gamma_{ist}$  – предсказанное значение моделью.

В отличие от среднеквадратических ошибок, где используется квадрат разности, MAE является линейной оценкой, поэтому вес разностей одинаков независимо от диапазона.

На рисунке 5 представлены графики динамики MSE, MAE согласно формулам (4), (5), которые позволяют оценить стабильность процесса обучения и соотношение ошибок на обучающей и валидационной (контрольной) выборках.

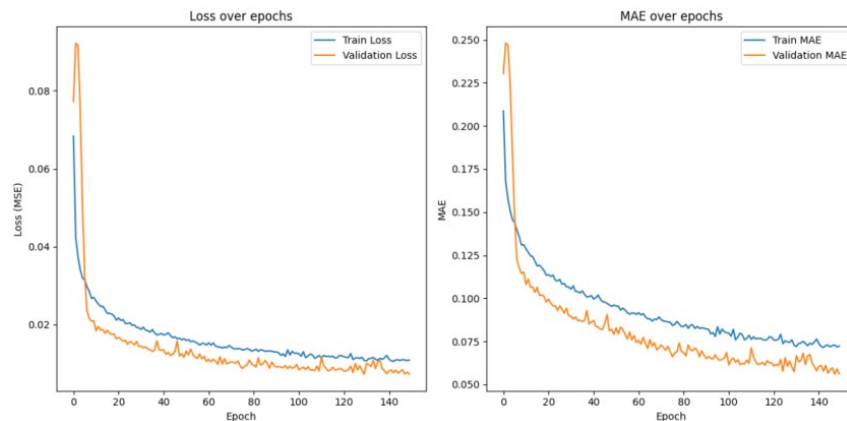


Рисунок 5 – Графики динамики MSE, MAE [материалы авторов]

### Практическая реализация и результаты

На рисунке 6 представлена диаграмма географического распределения вероятности, которая служит итоговым продуктом работы нейронной сети.

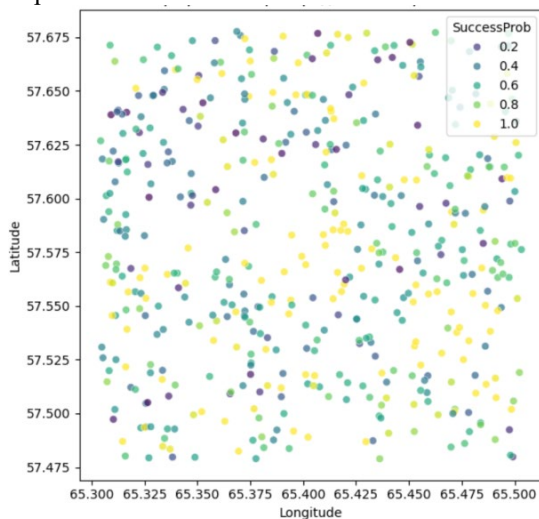


Рисунок 6 – Географическое распределение вероятности [материалы авторов]

На ней отображено визуальное распределение координат точек бурения, где каждая координата окрашена в свой цвет, что говорит о вероятности успеха бурения. На основании географического распределения на ограниченном объеме выборки была разработана интерактивная карта и пользовательский интерфейс для удобства взаимодействия системы и пользователя. На рисунке 7 представлен фрагмент интерактивной карты. На рисунке 8 представлен фрагмент интерактивной карты вблизи. В левом углу карты находится окно градации вероятности успеха добычи в конкретной точке. Основываясь на метриках гистограмм ошибок и графика остатков, можно сделать вывод, что модель имеет небольшие погрешности. Гистограмма ошибок показывает, что погрешности распределены равномерно, что может говорить о стабильности работы модели. График остатков показывает, что отклонение от действительных значений минимально. Несмотря на показатели метрик, для более точной и строгой оценки работы модели необходимо делать дополнительные тесты, в том числе на основе новых исходных данных.



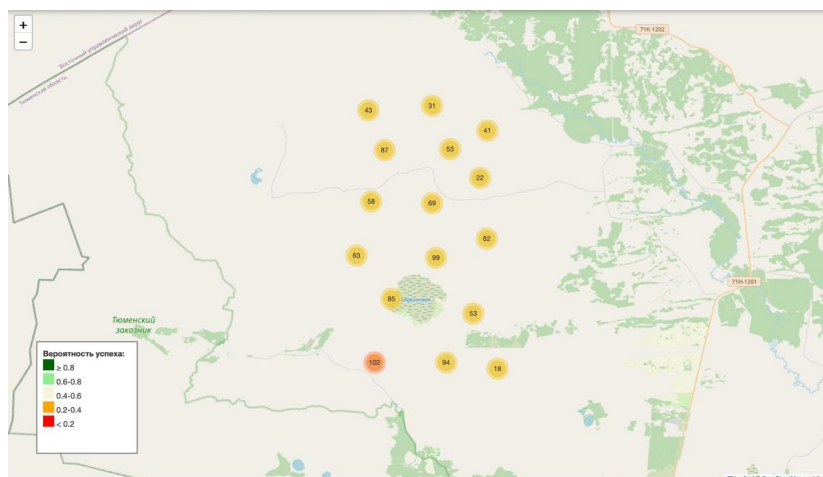


Рисунок 7 – Фрагмент интерактивной карты [материалы авторов]



Рисунок 8 – Приблизженный фрагмент интерактивной карты [материалы авторов]

Чем больше будет объем исходных данных, тем больше вероятность научить модель работать корректно, без смещений.

Важным этапом работы модели является анализ географического распределения вероятности успешности добычи нефти. Благодаря визуализации, пользователь может проанализировать более подробно местность, которая учувствует в исследовании. Участки на карте находятся друг от друга не небольшом расстоянии, а это значит, что в условиях реальных исследований, модель может облегчить задачу инженерам в выполнении анализа местности.

### Заклучение

На основе анализа распределений и визуализации зависимостей были выявлены следующие зависимости:

- при низкой пористости (5%) вероятность успешной добычи резко снижается. Такие породы плохо проницаемы для нефти;
- при средней пористости (10 – 20%) вероятность успеха значительно возрастает, модель воспринимает такие значения как оптимальные;

– при высокой пористости (20%) вероятность также остаётся высокой, но может быть ограничена другими параметрами (например, слишком низким давлением);

– глубины от 1000 до 4000 м являются наиболее благоприятными – на этих уровнях залегают основные продуктивные пласты;

– в оптимальном диапазоне (30 – 70 бар) наблюдается пик успешности – такие условия обеспечивают хорошее движение флюидов.

– чрезмерное давление (70 бар) может оказывать отрицательный эффект, особенно если сопровождается неблагоприятными геологическими условиями.

Для повышения прогностических способностей модели в будущем целесообразно расширить перечень входных параметров и обучить модель на других видах алгоритмов машинного обучения.

### Литература

1. Байбаров, Д. А. Оценка продуктивности и экономической эффективности технологий искусственного интеллекта для автоматизации

процессов разведки и добычи нефти и газа / Д. А. Байбаров // XXI век: итоги прошлого и проблемы настоящего плюс. – 2021. – Т. 10, № 3(55). – С. 100-105. – DOI 10.46548/21vek-2021-1055-0019.

2. Вершинин, В. Е. Нейросетевое моделирование Прогнозирование показателей добычи скважин в условиях нестационарного заводнения / В. Е. Вершинин, Р. Ю. Пономарев // Деловой журнал Neftegaz.RU. – 2022. – № 5-6(125-126). – С. 26-32. – EDN ZGEQVB.

3. Нейросетевое прогнозирование входных параметров при добыче нефти / Е. Д. Семенов, М. Я. Брагинский, Д. В. Тараканов, И. Л. Назарова // Вестник кибернетики. – 2023. – Т. 22, № 4. – С. 42-51. – DOI 10.35266/1999-7604-2023-4-6.

4. Патент № 2794707 C1 Российская Федерация, МПК E21B 41/00, E21B 49/00, G01V 99/00. Способ прогнозирования объемов добычи углеводородов из месторождений нефти и газа с использованием компьютерного моделирования: № 2022120987 : заявл. 02.08.2022 : опубл. 24.04.2023 / А. В. Бочкарев, Б. В. Васекин, Н. А. Воробьев [и др.] ; заявитель Общество с ограниченной ответственностью "Страта Солюшенс", Общество с ограниченной ответственностью "Физтех Геосервис".

5. Современные методы применения машинного обучения как инструмента прогнозирования добычи нефти / А. Р. Рустамов, Г. М. Пеньков, Д. Г. Петраков, М. А. Рустамова // Недропользование. – 2024. – Т. 24, № 1. – С. 44-50. – DOI 10.15593/2712-8008/2024.1.6.

6. Standards for Selection of Surfactant Compositions used in Completion and Stimulation Fluids / D. G. Petrakov, A. V. Loseva, N. T.

Alikhanov, H. Jafarpour // International Journal of Engineering. – 2023. – Vol. 36, No. 9. – P. 1605-1610. – DOI 10.5829/ije.2023.36.09c.03.

7. Цыпленков, С. В. Нейросетевой подход к ранжированию факторов, влияющих на энергоэффективность добычи нефти / С. В. Цыпленков, Е. Д. Агафонов, Д. И. Цыпленкова // Интеллектуальные системы в производстве. – 2022. – Т. 20, № 1. – С. 22-28. – DOI 10.22213/2410-9304-2022-1-22-28. – EDN KVZQAT.

8. Development of Monitoring and Forecasting Technology Energy Efficiency of Well Drilling Using Mechanical Specific Energy / A. Kunshin, M. Dvoynikov, E. Timashev, V. Starikov // Energies. – 2022. – Vol. 15, No. 19. – P. 7408. – DOI 10.3390/en15197408.

9. Ilyushin, Y.V., Nosova, V.A. (2025). Development of Mathematical Model for Forecasting the Production Rate. International Journal of Engineering, Transactions B: Applications, 38(8), 1749-1757. <https://doi.org/10.5829/ije.2025.38.08b.02>

10. Negash, B. M. Artificial neural network-based production forecasting for a hydrocarbon reservoir under water injection / B. M. Negash, A. D. Yaw // Petroleum Exploration and Development. – 2020. – Vol. 47, No. 2. – P. 383-392. – DOI 10.1016/s1876-3804(20)60055-6.

11. Колсун, Н. В. Внедрение искусственного интеллекта в горнодобывающем секторе / Н. В. Колсун, А. В. Гурко // Анализ и прогнозирование систем управления в промышленности, на транспорте и в логистике, Санкт-Петербург, 23–25 апреля 2024 года. – СПб.: ООО "Медиапапир", 2024. – С. 223-229.

**Сорокин А.А., Гурко А.В. Исследование методов прогнозирования добычи нефти.** В статье анализируются методы прогнозирования добычи нефти с применением нейронных сетей, выполняется компьютерный эксперимент оценки вероятности успешной добычи нефти на основе анализа геофизических параметров. Представлена архитектура нейронной модели, приведены результаты её обучения и тестирования. Направлением дальнейших исследований является расширение перечня входных параметров и обучение модели на других видах алгоритмов машинного обучения для повышения прогностических способностей модели.

**Ключевые слова:** интеллектуальная система, нейронные сети, машинное обучение, прогнозирование добычи нефти, нефтяная отрасль, глубокое обучение, автоматизация, моделирование, искусственный интеллект.

**Sorokin A.A., Gurko A.V. Methods of research of oil production forecasting.** The article considers modern forecasting methods in the oil industry, the use of neural networks in forecasting oil production, and the development of an application solution for automating the assessment of the probability of successful oil production based on the analysis of geophysical parameters. The architecture of the developed model is introduced, and the results of its training and testing are given.

**Keywords:** intelligent system, neural networks, machine learning, oil production forecasting, oil industry, deep learning, automation, modeling, artificial intelligence.

Статья поступила в редакцию 15.05.2025  
Рекомендована к публикации профессором Мальчевой Р. В.