

УДК 622.276.53:621.3
DOI: 10.18799/24131830/2025/1/4610
Шифр специальности ВАК: 2.4.2

Диагностика неисправностей штанговой глубинно-насосной установки на основе машинного обучения с использованием кривой мощности двигателя

О.Х. Ахмед¹, С.И. Текле², А.М. Зюзев¹, В.П. Метельков¹✉

¹ Уральский федеральный университет, Россия, г. Екатеринбург

² Колледж инженерии и технологий, Эрпретя, Май Нефти

✉ v.p.metelkov@urfu.ru

Аннотация. Актуальность. Определяется сложностью мониторинга и диагностики состояния подземных элементов конструкции штанговых глубинно-насосных установок и больших экономических потерь при эксплуатации этого оборудования с не выявленными своевременно дефектами. **Цель:** разработка способов обнаружения неисправностей штангового насоса, не требующих привлечения для диагностики высококвалифицированного персонала, с использованием информации, получение которой легко доступно на поверхности. **Методы:** методы машинного обучения (метод дерева решений, метод К-ближайших соседей, метод опорных векторов, наивный байесовский классификатор) с использованием диаграмм мощности двигателя. **Результаты и выводы.** Показана возможность обнаружения неисправностей штангового насоса на основе методов машинного обучения. Исследование проведено на основе разработанной имитационной модели штангового насоса, использованной для воспроизведения кривых мощности двигателя с учетом влияния особенностей различных сценариев функционирования оборудования. Являясь основным источником энергии для добычи нефти, мощность двигателя напрямую связана с рабочим состоянием нефтяной скважины в реальном времени, а кривая мощности двигателя является надежным источником информации, позволяющим повысить эффективность диагностики штангового насоса. Для обучения классификаторов и оценки точности их функционирования использовался ряд характеристик, получаемых из кривых мощности двигателя для шести различных рабочих состояний насоса. А именно, рассчитывались рабочие коэффициенты, представляющие собой отношения интеграла мощности на каждом из четырех этапов цикла работы установки к интегралу мощности за весь цикл. Результаты показывают, что рассматриваемый подход позволяет обеспечить высокую точность диагностики условий работы штангового насоса. Классификатор на основе метода решающих деревьев показал наивысшую эффективность среди четырех исследованных классификаторов по выявлению всех шести типов неисправностей (95,8 %), метод опорных векторов также показал весьма высокую эффективность (90,3 %).

Keywords: штанговая глубинно-насосная установка, кривая мощности двигателя, динамограмма, диагностика неисправностей, алгоритмы машинного обучения, метод дерева решений, метод К-ближайших соседей, метод опорных векторов, наивный байесовский классификатор

Для цитирования: Диагностика неисправностей штанговой глубинно-насосной установки на основе машинного обучения с использованием кривой мощности двигателя / О.Х. Ахмед, С.И. Текле, А.М. Зюзев, В.П. Метельков // Известия томского политехнического университета. Инжиниринг георесурсов. – 2025. – Т. 336. – № 1. – С. 36–49. DOI: 10.18799/24131830/2025/1/4610

UDC 622.276.53:621.3

DOI: 10.18799/24131830/2025/1/4610

Machine learning based sucker rod pump fault diagnosis using motor power curve

O.H. Ahmed¹, S.I. Tecle², A.M. Zyuzev¹, V.P. Metelkov¹✉

¹ Ural Federal University, Ekaterinburg, Russian Federation

² College of Engineering and Technology, Mai Nefhi, Eritrea

✉ v.p.metelkov@urfu.ru

Abstract. Relevance. The complexity of monitoring and diagnosing the condition of underground structural elements of sucker rod pumping units and large economic losses when operating this equipment with defects not identified in a timely manner. **Aim.** Development of methods for detecting faults in a sucker rod pump that do not require the involvement of highly qualified personnel for diagnosis, using information that is easily available on the surface. **Methods.** Machine learning methods (Decision tree method, K-nearest neighbors method, Support vector machine, Naive Bayes classifier) using motor power curves. **Results and conclusions.** The paper demonstrates the possibility of detecting faults in a sucker rod pump based on machine learning methods. The study was carried out on the basis of a developed simulation model of a sucker rod pump, used to reproduce motor power curves, taking into account the impact of the features of various equipment operation scenarios. Being the fundamental energy source for the oil production, motor power is directly related to the real-time operating condition of the oil well, and the motor power curve is a reliable source with the ability to increase the efficiency of sucker rod pump diagnostic. To train the machine learning classifiers and evaluate their performance accuracy, a number of characteristics were used, obtained from motor power curves for six different pump operating states. Namely, operating coefficients were calculated, representing the ratio of the power integral at each of the four stages of the installation operating cycle to the power integral for the entire cycle. The results show that the considered approach allows for high accuracy in diagnosing the operating conditions of a sucker rod pump. The classifier based on the decision tree method showed the highest efficiency among the four studied classifiers in identifying all six types of faults (95.8%), and the support vector machine method showed as well very high efficiency (90.3%).

Keywords: sucker rod pump, motor power curve, dynamometer card, fault diagnosis, machine learning algorithms, Decision tree, K-nearest Neighbors, SVM, Naive Bayes classifier

For citation: Ahmed O.H., Tecle S.I., Zyuzev A.M., Metelkov V.P. Machine learning based sucker rod pump fault diagnosis using motor power curve. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2025, vol. 336, no. 1, pp. 36–49. DOI: 10.18799/24131830/2025/1/4610

Введение

Штанговые глубинно-насосные установки (ШГНУ) широко используются в нефтедобывающей отрасли, поскольку они имеют простое оборудование, низкую совокупную стоимость и подходят для широкого диапазона производительности скважин. В настоящее время около 80 % нефтяных скважин в мире используют технологию добычи с использованием штанговых насосов [1]. Производительность и затраты на эксплуатацию штангового насоса зависят от условий его работы, а из-за разнообразных и сложных условий эксплуатации системы штанговых насосов зачастую работают в нештатных условиях, что приводит к увеличению частоты отказов оборудования или снижению его производительности. В литературе наряду с нормальным функциональным состоянием выделяют более девяти проблемных рабочих состояний. Поэтому крайне важно использовать надежную и точ-

ную систему диагностики неисправностей, чтобы поддерживать функционирование штангового насоса в нормальном рабочем состоянии.

Как правило, системы обнаружения неисправностей базировались на использовании динамограмм (замкнутом графике смещения полированного штока в зависимости от нагрузки на полированный шток) в качестве основного источника данных. Различные состояния функционирования штангового насоса четко идентифицируются по форме динамограммы [2]. Для получения динамограммы можно использовать контроллер штангового насоса, оснащенный встроенными датчиками нагрузки и положения. Эти датчики положения обеспечивают измерение перемещения полированного штока с помощью различных средств, таких как позиционный переключатель, инклинометр или преобразователь на эффекте Холла. Обычно требуется достаточно высокая квалификация персонала для полу-

чения достоверных результатов путем исследования динамограммы. При этом принятие решений с использованием такого подхода весьма субъективно и не позволяет проводить мониторинг в реальном времени. Чтобы решить эту проблему, в работах [3–7] предлагаются некоторые весьма сложные методы для извлечения необходимых признаков с последующей классификацией.

Колонна насосных штанг в насосной системе может испытывать статические и динамические силы, связанные с вынужденными и свободными колебаниями, эффекты этих сил передаются на приводной двигатель [7]. Следовательно, изменения нагрузки на плунжер можно наблюдать по измеренным сигналам на входных клеммах двигателя. Примечательно, что электрические измерения, особенно диаграмма мощности двигателя, весьма точно отражают состояние пласта. Диаграмма мощности двигателя является периодической, соответствующей продолжительности одного хода, и эффективно отражает влияние изменения нагрузки с течением времени. Такие электрические измерения не ставят под угрозу надежность системы и обеспечивают достаточно высокую точность. Можно отметить работы [8, 9], исследователи использовали диаграмму мощности двигателя в качестве альтернативного источника данных в системе диагностики неисправностей штанговых насосов. Возможности современных интеллектуальных станций управления ШГНУ позволяют реализовать функции диагностики с использованием динамограмм и графиков мощности двигателя [10].

Методы обнаружения неисправностей штанговых насосов по данным о графике мощности двигателя можно разделить на два подхода – прямой и косвенный. При использовании косвенного подхода нагрузку на полированный шток можно определить на основе электрических измерений с использованием метода коэффициента крутящего момента или с помощью модели на основе одномерного уравнения, описывающего волновые явления с учетом различных подземных параметров и измерений [11]. Однако измерение этих параметров ниже поверхности земли оказывается затруднительным. Кроме того, небольшая величина коэффициента крутящего момента вблизи нижней и верхней мертвых точек приводит к снижению точности получаемой динамограммы. При прямом подходе на основе кривых мощности двигателя имеется возможность извлечения характерных особенностей, которые четко представляют каждое рабочее состояние, позволяя обойти отмеченные выше недостатки.

Проблема классификации – другой вопрос, который необходимо решить при разработке диагностической модели. Задача состоит в соотношении

входного вектора признаков одному из K рабочих условий, желательно в режиме реального времени. Эта задача сводится к проблеме распознавания образов, которая в современной практике решается с использованием различных методов машинного обучения (Machine Learning, ML), таких как Искусственная нейронная сеть (Artificial Neural Network, ANN), Скрытая марковская модель (Hidden Markov Model, HMM), Метод K -ближайших соседей (K-Nearest Neighbors, KNN), Метод решающего дерева (Decision Tree, DT), Метод опорных векторов (Support Vector Machine, SVM), Экстремальная обучающая машина (Extreme Learning Machine, ELM) и другие. HMM – это модель временных рядов, которая использует последовательные шаблоны данных, в то время как другие методы, такие как DT, SVM, KNN, наивный байесовский метод (Naive Bayes classifier, NB), ANN и ELM, представляют собой диагностическую модель, основанную на предположении, что наборы точек данных независимы и одинаково распределены. Методы KNN в [12], SVM в [3, 12], ELM в [13], ANN в [4, 14] и HMM в [9] использовались для классификации случаев неисправности штанговых насосов. В отличие от HMM, методы, основанные на SVM, ELM и ANN, не требуют обязательного описания процесса, поскольку они просто оценивают входные и выходные данные для разработки диагностических моделей.

Еще одна проблема связана с получением достаточного количества обучающих выборок, представляющих различные условия эксплуатации, для создания полного обучающего набора, необходимого для методов машинного обучения. Важно не допустить заблуждения, что эти выборки независимы и одинаково распределены. Корректный подход предполагает ослабление требований к независимости выборки и идентичности распределения. В этих обстоятельствах кажется, что предпочтительным методом машинного обучения для решения проблемы классификации является скрытая марковская модель. Однако появление отдельных эксплуатационных состояний ограничено на протяжении всего жизненного цикла скважины (обычно менее четырех). Следовательно, становится непрактичным аппроксимировать настоящую апостериорную плотность вероятности, используя исторические данные. Можно отметить, что это ограничение не распространяется на алгоритмы машинного обучения, которые используются в данном исследовании, такие как метод дерева решений, метод K -ближайших соседей, метод опорных векторов и наивный Байесовский метод, из-за отсутствия в них апостериорных вероятностей. В результате эти алгоритмы могут достаточно эффективно использовать данные, сгенерированные моделью.

Кроме того, эти алгоритмы обладают хорошей способностью использовать небольшие наборы данных для решения задач классификации и распознавания [15]. Таким образом, в данном исследовании, по сравнению с более ранней работой [16], сфера применения алгоритмов машинного обучения расширена за пределы использования только метода опорных векторов и включает в себя другие три метода машинного обучения, которые упомянуты выше. В работе [16] были проанализированы шесть различных сценариев эксплуатации штанговых насосов, но диагностический анализ был ограничен только четырьмя вариантами неисправностей. В данном исследовании классифицируются все шесть вариантов на основе данных о мощности двигателя.

Анализ работы штангового насоса

Конфигурация ШГНУ включает в себя приводной двигатель, редуктор, поверхностное оборудование, колонну насосных штанг и насосно-компрессорные трубы. В большинстве штанговых насосных установок основным источником энергии служат электродвигатели, работающие на высокой скорости, которая впоследствии снижается редуктором в соотношении примерно 30:1 [17]. В насосной установке используется четырехрычажный приводной механизм для преобразования вращательного движения в возвратно-поступательное. Это возвратно-поступательное движение затем передается глубинному насосу через колонну насосных штанг, обеспечивая подъем нефти из скважины на поверхность.

Всасывающий клапан и нагнетательный клапан насоса включаются в работу при возвратно-поступательном движении полированного штока. Хотя открытие обоих клапанов зависит и от количества свободного газа в насосе, в нормальном рабочем состоянии всасывающий клапан закрывается в конце хода вверх, а нагнетательный клапан закрывается в конце хода вниз. От состояния этих клапанов зависит давление насоса, определяющее нагрузку на плунжер. Когда нагнетательный клапан открыт, давление насоса почти равно давлению нагнетания, и на плунжер действует только выталкивающая сила. В то время как при открытом всасывающем клапане давление насоса практически равно давлению всасывания и на плунжер оказывается большая нагрузка. Неравномерная нагрузка на плунжер во время цикла приводит к неравномерной нагрузке редуктора. Поэтому в ШГНУ на кривошипе закреплен противовес для компенсации неравномерности нагрузок и уменьшения мощности двигателя.

При ходе вверх противовес перемещается от положения с большей потенциальной энергией к положению с более низкой, а при ходе вниз потенциальная энергия противовеса постепенно увеличива-

ется. В результате противовес помогает двигателю поднимать тяжелый груз во время хода вверх. И наоборот, при ходе вниз вес колонны насосных штанг помогает двигателю поднимать противовес.

На рис. 1 поясняется механизм работы штангового насоса. В нормальном рабочем состоянии насоса динамограмма по форме приближается к параллелограмму, а кривая изменения мощности двигателя примерно соответствует синусоидальной кривой. Идентификация пиков и спадов на кривой мощности двигателя, а также четырех углов динамограммы оказывается весьма удобной для анализа рабочего процесса. Путь ABC (рис. 1) соответствует ходу вверх, при котором всасывающий клапан открывается в точке В и закрывается в точке С. Путь CDA соответствует ходу вниз, при этом нагнетательный клапан открывается в точке D и закрывается в точке А.

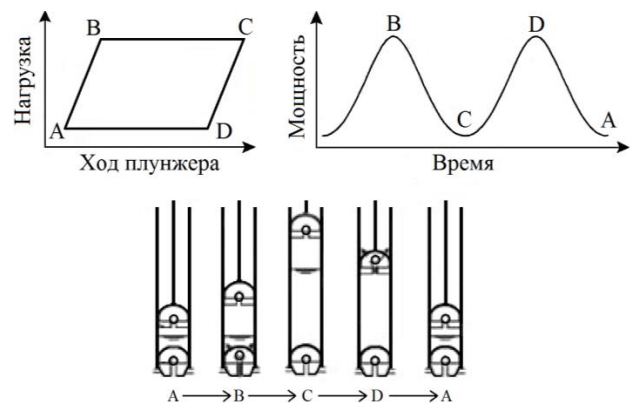


Рис. 1. Диаграммы для анализа механизма ШГНУ

Fig. 1. Diagram for analysis of sucker rod pump mechanism

Система диагностики неисправностей штангового насоса

Изменения условий работы в подземной части установки изменяют нагрузочную кривую двигателя, что отражается на формах динамограммы и кривых мощности двигателя. Но их идентификация даже опытными инженерами посредством интерпретации динамограмм неэффективна из-за ее субъективного характера и необходимости большого труда при подготовке обучающих выборок [18]. Экспериментальные методы, используемые для получения динамограмм, требуют существенных затрат на техническое обеспечение, а использование мощности двигателя в качестве альтернативы для диагностических моделей не столь требовательно к этим затратам. Однако визуальные методы оценки по кривым мощности весьма затруднительны, что подчеркивает необходимость компьютерных методов для эффективного использования этого подхода.

В данном исследовании, помимо метода опорных векторов, для классификации неисправностей штанговых насосов были использованы еще три метода машинного обучения, относящиеся к семейству алгоритмов обучения с учителем, с целью сравнительной оценки их точности и пригодности для выявления такого рода неисправностей. Этими методами являются: метод дерева решений, метод К-ближайших соседей, метод опорных векторов и наивный байесовский классификатор. Системы мониторинга состояния, основанные на машинном обучении, работают с использованием исторических данных, полученных от датчиков на механизме в различных условиях эксплуатации, которые впоследствии используются для извлечения признаков с целью обучения классификаторов. На завершающем этапе модели классификаторов проходят процедуру перекрестной проверки, которая служит для оценки эффективности алгоритма машинного обучения при его работе с новыми наборами данных, на которых классификатор не обучался.

Извлечение признаков

На точность работы диагностической модели сильно влияет построение этапа извлечения признаков. В исследовании были использованы признаки, определяющие характеристики кривой мощности двигателя, полученные на основе положений рабочих точек клапана и относительной работы, совершаемой на разных участках хода. Использовались семь специфических признаков состояния ($W_1, W_2, W_3, W_4, L_a, L_b, L_c$):

Коэффициенты работы для каждой четырех частей цикла:

$$W_1 = \frac{1}{W} \int_{t_0}^{t_1} p(t) dt; \quad (1)$$

$$W_2 = \frac{1}{W} \int_{t_1}^{t_2} p(t) dt; \quad (2)$$

$$W_3 = \frac{1}{W} \int_{t_2}^{t_3} p(t) dt; \quad (3)$$

$$W_4 = \frac{1}{W} \int_{t_3}^{t_4} p(t) dt, \quad (4)$$

где

$$W = \int_{t_0}^{t_4} p(t) dt. \quad (5)$$

- L_a – расстояние между положениями, когда открыт всасывающий клапан и нагнетательный:

$$L_a = \sqrt{(x_{SVo} - x_{TVo})^2 + (y_{SVo} - y_{TVo})^2}. \quad (6)$$

- L_b – расстояние между положениями, когда закрыт и открыт всасывающий клапан:

$$L_b = \sqrt{(x_{SVc} - x_{SVo})^2 + (y_{SVc} - y_{SVo})^2}. \quad (7)$$

- L_c – расстояние между положениями, когда закрыт и открыт нагнетательный клапан:

$$L_c = \sqrt{(x_{TVc} - x_{TVo})^2 + (y_{TVc} - y_{TVo})^2}. \quad (8)$$

В выражениях (1)–(5) обозначено: p – мгновенная мощность двигателя; t_0-t_1 – продолжительность первой части хода вниз; t_1-t_2 – продолжительность второй части хода вниз; t_2-t_3 – продолжительность первой части хода вверх; t_3-t_4 – продолжительность второй части хода вверх. В выражениях (6)–(8) обозначено: x_i – положение i -ой точки на графике мощности по горизонтальной оси (оси времени) в относительных единицах, где за базисное время принято время цикла; y_i – положение i -ой точки на графике мощности по вертикальной оси (оси мощности) в относительных единицах, где за базисную мощность принята номинальная мощность двигателя. Индексы i в выражениях (6)–(8): $i=SVo$ – открыт всасывающий клапан; $i=TVo$ – открыт нагнетательный клапан; $i=SVc$ – закрыт всасывающий клапан; $i=TVc$ – закрыт нагнетательный клапан.

Метод дерева решений

Дерево решений – это непараметрический метод обучения, который служит диагностическим инструментом, создающим систему, основанную на знаниях, посредством индуктивного вывода из историй прецедентов. Дерево начинается с корневого узла, разделяющегося на две ветви; дальнейшее разделение происходит на промежуточных узлах, также называемых узлами принятия решений. Завершается дерево конечными или листовыми узлами, и любое связанное подмножество узлов называется поддеревом (рис. 2).

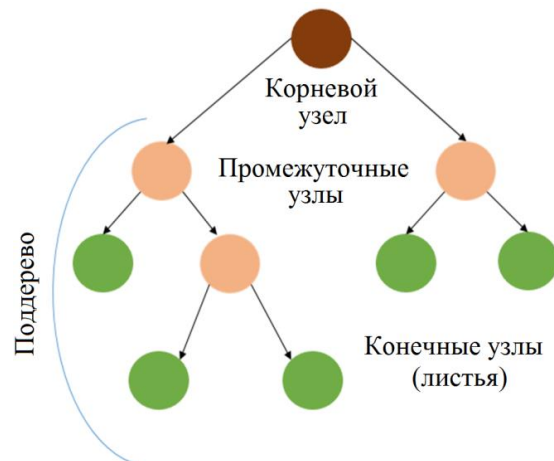


Рис. 2. Структура дерева решений

Fig. 2. Decision tree structure

Структура дерева решений опирается на выбор корня, определяемый теорией информации Куинлана (прирост информации) [19]. Согласно этому критерию, информация в сообщении зависит от ее вероятности [20]. Дерево решений строится на ос-

нове обучающего набора T , состоящего из случаев, задающих значения атрибутов и классов, обозначаемых $\{C_1, C_2, \dots, C_k\}$. Если существует тест с n результатами, разделяющий T на подмножества T_1, T_2, \dots, T_n , и S – любой набор случаев, $freq(C_i, S)$ – это количество случаев в S , принадлежащих классу C_i , и $|S|$ – общее количество случаев в наборе S . Если случай выбран из S случайным образом и принадлежит классу C_j , сообщение имеет вероятность $freq(C_j, S)/|S|$, и информация, ему соответствующая, составляет $-\log_2(freq(C_j, S)/|S|)$ битов.

Ожидаемая информация, необходимая для определения класса случая в S :

$$info(S) = - \sum_{j=1}^k \frac{freq(C_j, S)}{|S|} * \log_2 \left(\frac{freq(C_j, S)}{|S|} \right).$$

Применительно к ситуации обучения $info(T)$ оценивает среднюю информацию, необходимую для определения класса случая в пределах T . Аналогично после того, как T было разделено в соответствии с n результатами теста X :

$$info_x(T) = - \sum_{i=1}^n (|T_i|/|T|) * info(T_i).$$

Лучшим атрибутом разделения будет тот, который обеспечит максимальный прирост информации результирующего узла относительно исходного. Прирост информации составляет:

$$Gain(X) = info(T) - info_x(T).$$

Метод К-ближайших соседей

Метод К-ближайших соседей представляет собой классификатор, который, несмотря на свою простоту, демонстрирует высокую эффективность. Важными для реализации этого классификатора параметрами являются целое положительное число (K) и метрика (d). Классификация данных тестирования определяется метками K ближайших к точке соседей [21]. Рис. 3 иллюстрирует эту концепцию: когда $K=1$, новая точка относится к Классу 1, поскольку в Классе 1 имеется только одно наблюдение, а в Классе 2 нет наблюдений; когда $K=2$, новая точка относится к классу 2, поскольку большинство наблюдений в этой области относится к Классу 2.

В методе KNN требуется определение расстояния между точкой входных данных и другими точками из набора обучающих данных. Для двух чек x и y с координатами (x_1, x_2, \dots, x_n) и (y_1, y_2, \dots, y_n) расстояние между ними вычисляется следующим образом:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2},$$

где x, y – две точки в n -мерном евклидовом пространстве.

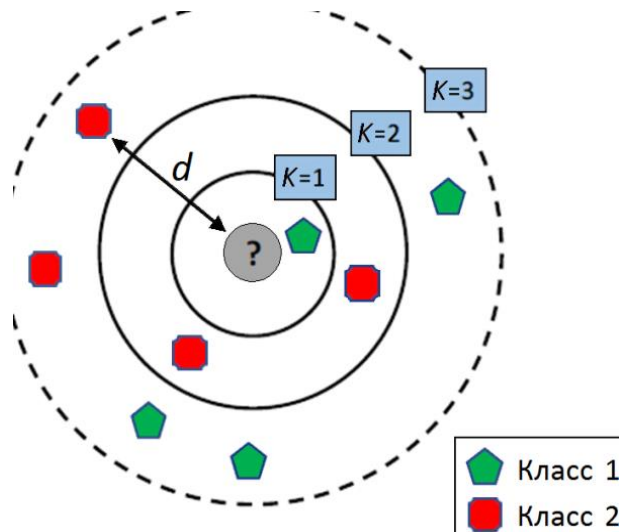


Рис. 3. Иллюстрация алгоритма KNN [22]

Fig. 3. KNN algorithm illustration [22]

Метод опорных векторов

Метод опорных векторов – это метод вычислительного обучения для классификации небольших выборок. Этот метод разделяет данные на положительные и отрицательные классы, определяя некоторую оптимальную гиперплоскость, которая делит все точки данных на эти две группы. Уравнение гиперплоскости можно записать как

$$y(x) = W^T x + b = 0,$$

где x – вектор признаков; W – n -мерный весовой вектор, нормальный к гиперплоскости; b – скалярный параметр-смещение. Эти параметры используются для нахождения наилучшей гиперплоскости – гиперплоскости с наибольшим зазором между двумя классами. Ближайшие точки, определяющие зазор между классами, называются опорными векторами. Если вектор признаков удовлетворяет условию $y(x_i) = W^T x_i + b \geq 1$, он принадлежит к классу А. Если $y(x_i) = W^T x_i + b \leq -1$ – к классу В.

Наивный байесовский классификатор

Наивный байесовский классификатор представляет собой алгоритм классификации, работающий на основе теоремы Байеса. Классификатор строится с опорой на «наивное» предположение, что каждая пара признаков условно независима и зависит от значения переменной класса. Формально для переменной класса y и зависимых векторов признаков от x_1 до x_n теорема Байеса устанавливает следующие отношения:

$$P(y|x_i) = \frac{P(x_i|y)P(y)}{P(x_i)}, \quad (9)$$

где $x_i=(x_1, x_2, x_3, \dots, x_n)$.

В (9) обозначено: $P(y|x_i)$ – апостериорная вероятность, $P(x_i|y)$ – правдоподобие, $P(y)$ – априорная вероятность класса; $P(x_i)$ – априорная вероятность предиктора.

С учетом «наивного» предположения об условной независимости можно записать:

$$P(x_i|y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = P(x_i|y).$$

Тогда соотношение (9) упрощается до следующего:

$$P(y|x_i) = \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_i)}.$$

Поскольку априорная вероятность предиктора $P(x_i)$ постоянна с учетом входных данных, получаем следующее правило классификации:

$$P(y|x_i) \propto P(y) \prod_{i=1}^n P(x_i|y).$$

Это значит, что задача наивной байесовской классификации состоит в том, что для разных значений класса y надо найти максимум $P(y|x_i)$, пропорциональной $P(y) \prod_{i=1}^n P(x_i|y)$. Для определения априорной вероятности класса $P(y)$ можно использовать относительную частоту класса y в обучающих данных.

Результаты моделирования

Получение набора данных для обучения

Использование методов компьютерной диагностики повышает эффективность контроля за работой штанговых насосов. Однако эти методы требуют обширного набора данных для описания различных рабочих состояний [9], даже несмотря на то, что за срок службы скважины возникает ограниченное число таких состояний. Для решения этой проблемы создана имитационная модель, изображенная на рис. 4, которая включает в себя модель пласта, модель насоса, модель колонны насосных штанг, модель первичного двигателя и модель насосного узла. На рис. 4 обозначено: V_s –

напряжение питания; F_{phl} – нагрузка на полированный шток; h_{phl} – перемещение полированного штока; h_{pl} – перемещение плунжера; v_{pl} – скорость плунжера; F_{pl} – нагрузка на плунжер; P_{rev} – пластовое давление; P_{wf} – забойное давление; P_s – давление всасывания; P_d – давление нагнетания; q – скорость потока жидкости.

Блок колонны насосных штанг включает в себя модель колонны насосных штанг как механического звена с распределенными параметрами с учетом как поверхностных, так и подземных граничных условий. Блок, включающий в себя электродвигатель и насосную установку, рассчитывает кинематические параметры полированного штока. Модель насоса определяет нагрузку на плунжер. Корректировка параметров потока, свойств жидкости или принятие допущений о различных механических проблемах в модели пласта позволяет моделировать различные сценарии подземной среды. Блок балансировки рассчитывает забойное давление и давление всасывания на основе скорости притока и жидкости, поступающей в насос.

Исследуя возможные вариации в рамках заданного сценария под поверхностью земли, разработанная модель позволяет получить различные динамограммы и кривые мощности двигателя, которые можно использовать для построения обучающего набора. В этот набор входят образцы, представляющие нормальные рабочие состояния, недостаточную подачу жидкости, утечку нагнетательного клапана, воздействие газа, попадание плунжера в верхнюю мертвую точку, попадание плунжера в нижнюю мертвую точку.

В данном исследовании при построении диагностической модели используется только кривая мощности двигателя. На рис. 5–10 показаны типичные динамограммы (а) и соответствующие им кривые мощности двигателя (b) для различных рабочих состояний.

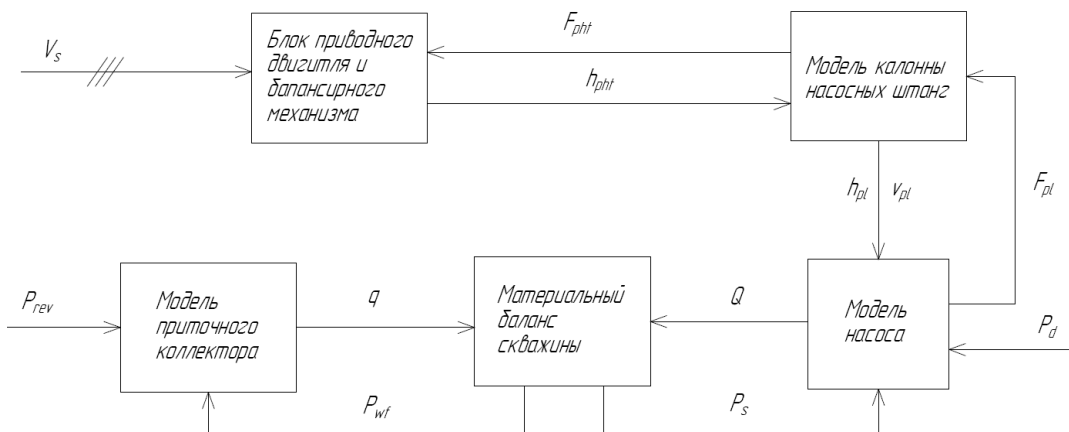


Рис. 4. Имитационная модель штангового насосного агрегата
 Fig. 4. Simulation model for sucker rod pump

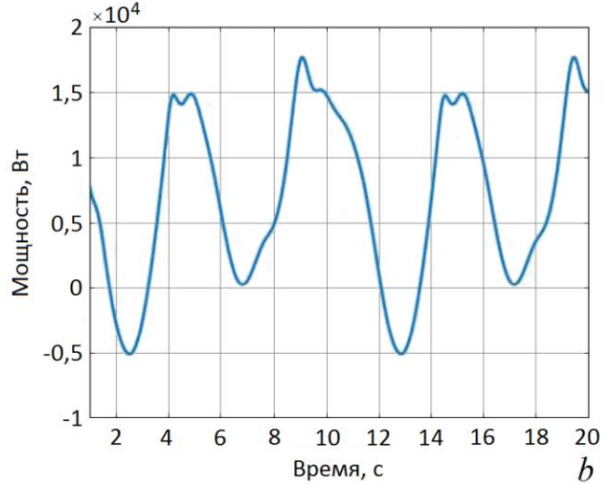
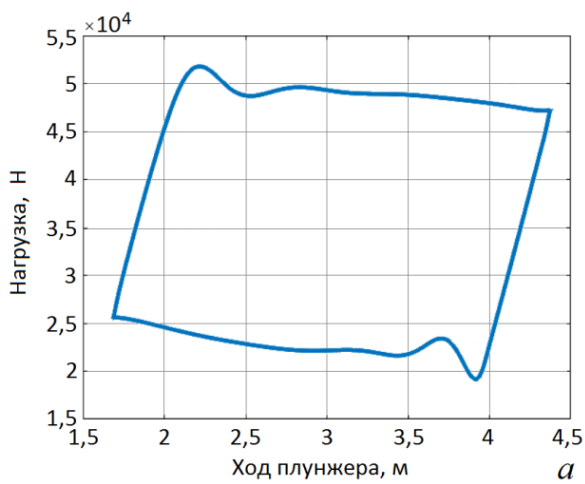


Рис. 5. *Нормальное рабочее состояние*
Fig. 5. *Normal working state*

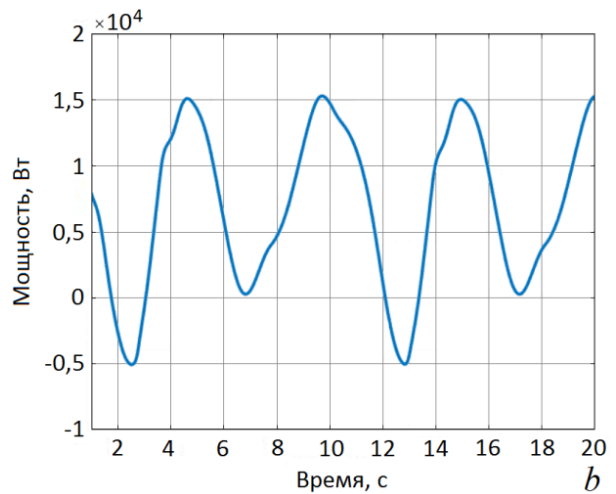
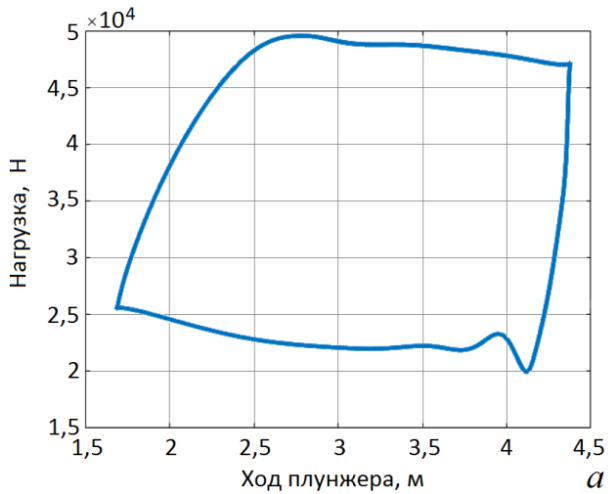


Рис. 6. *Негерметичность нагнетательного клапана*
Fig. 6. *Travelling valve leakage*

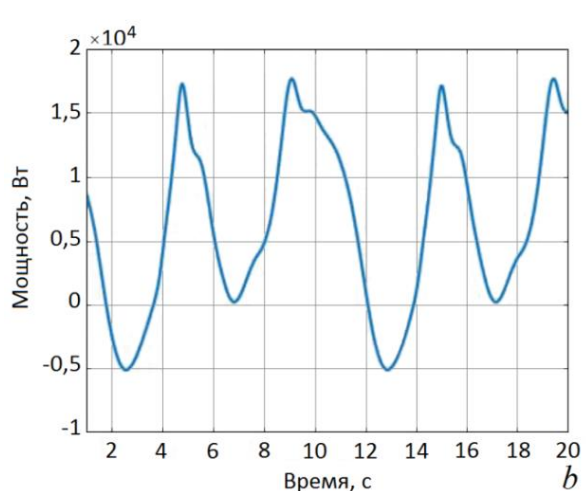
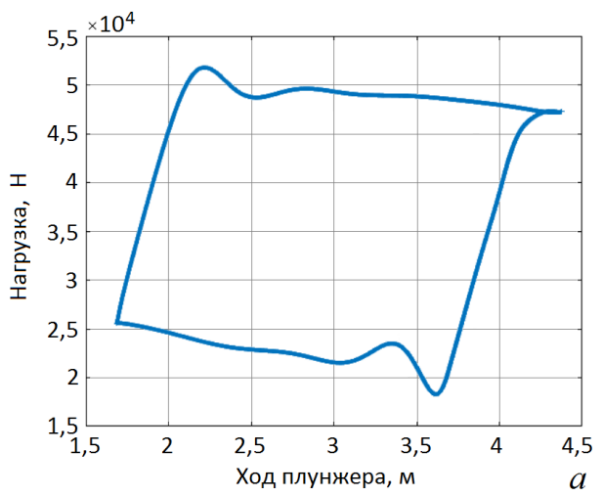


Рис. 7. *Эффект захвата газа*
Fig. 7. *Gas affected*

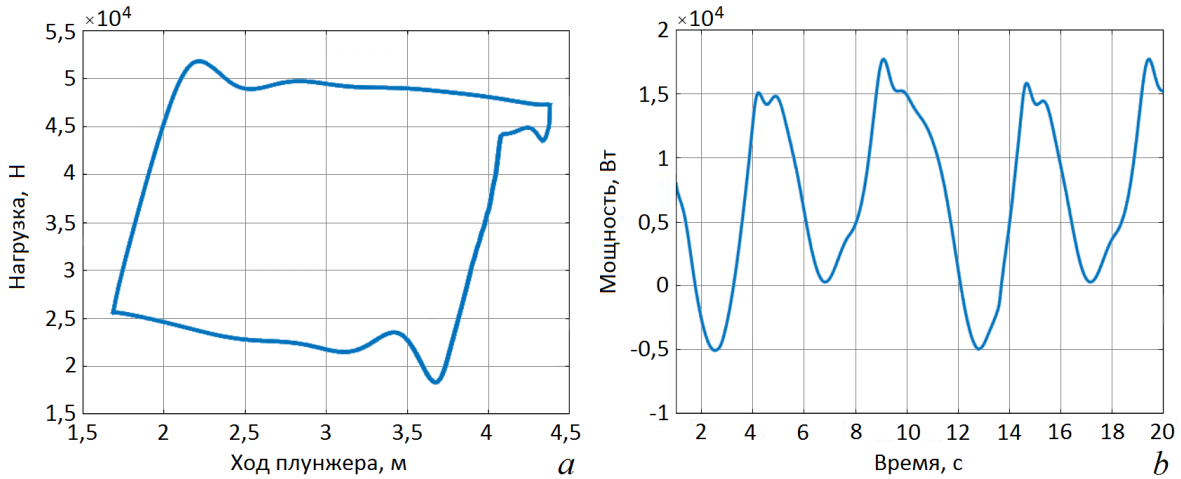


Рис. 8. Недостаточное поступление жидкости
Fig. 8. Insufficient liquid supply

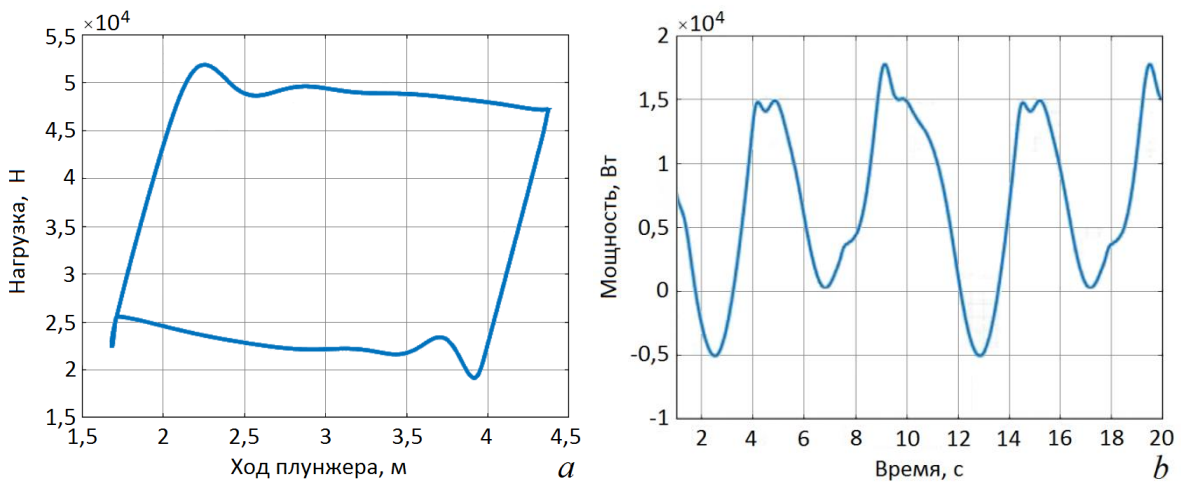


Рис. 9. Состояние, когда плунжер достигает нижней мертвой точки
Fig. 9. Condition when the plunger hits bottom dead center

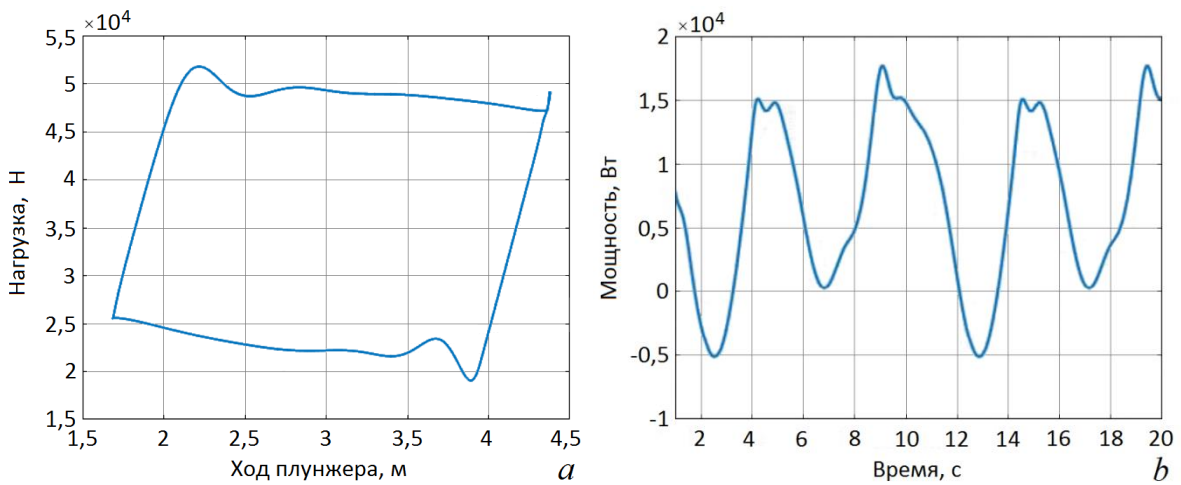


Рис. 10. Состояние, когда плунжер достигает верхней мертвой точки
Fig. 10. Condition when the plunger hits top dead center

Извлечение признаков

На рис. 11 показаны нормированные кривые смещения и скорости полированного штока, нагрузки на него, мощности двигателя и давления насоса. На кривой давления насоса легко различимы положения рабочих точек клапанов. Начало сжатия соответствует закрытию всасывающего клапана, а начало расширения соответствует закрытию нагнетательного клапана. Аналогично конец сжатия совпадает с открытием нагнетательного клапана, а конец расширения – с открытием всасывающего клапана. При сравнении кривой давления насоса с кривой мощности двигателя положения открытия, всасывающего и нагнетательного клапанов на кривой мощности двигателя совпадают с пиками, а положения закрытия клапанов соответствуют впадинам. Поэтому пики и провалы на кривой мощности двигателя имеют решающее значение для описания рабочего процесса. В рамках предложенного метода для каждого образца определяются векторы признаков, в таблице представлены средние значения компонентов этих векторов. Можно сделать вывод, что эти векторы признаков весьма отчетливо характеризуют каждое рабочее состояние.

Таблица. Средние значения извлекаемых признаков

Table. Average values of extracted feature

Состояние Condition	L_a	L_b	L_c	W_1	W_2	W_3	W_4
Nor	0,425	0,797	0,453	0,397	0,217	0,202	0,184
Lik	0,477	0,729	0,462	0,355	0,217	0,241	0,187
Gaf	0,416	0,799	0,503	0,448	0,247	0,102	0,204
Ins	0,398	0,793	0,450	0,486	0,257	0,054	0,203
Vpb	0,442	0,711	0,437	0,185	0,183	0,409	0,222
Ins	0,436	0,702	0,434	0,204	0,187	0,388	0,222

В таблице обозначено: Nor – нормальное рабочее состояние; Lik – негерметичность нагнетательного клапана; Gaf – эффект захвата газа; Ins – недостаточ-

ное поступление жидкости; Vpb – состояние, когда достигается нижняя мертвая точка, и Trpb – состояние, когда плунжер достигает верхней мертвой точки.

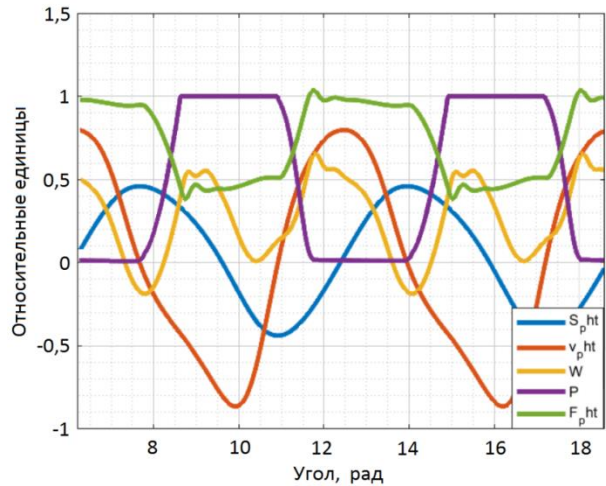


Рис. 11. Динамические характеристики механизма
Fig. 11. Mechanism dynamic characteristics

Классификация

На рис. 12 представлена схема диагностической модели на основе упомянутых выше алгоритмов машинного обучения, где процесс диагностики делится на обучение и собственно диагностику. На этапе обучения на основе кривых мощности двигателя создается учебный набор. Выбранные ML-алгоритмы обучаются с использованием этого набора. Обученные модели проходят проверку с помощью четырехкратной перекрестной проверки. В процессе диагностики кривая мощности двигателя обрабатывается в режиме реального времени для создания вектора признаков. Затем классификаторы используют этот вектор признаков для идентификации рабочего состояния.

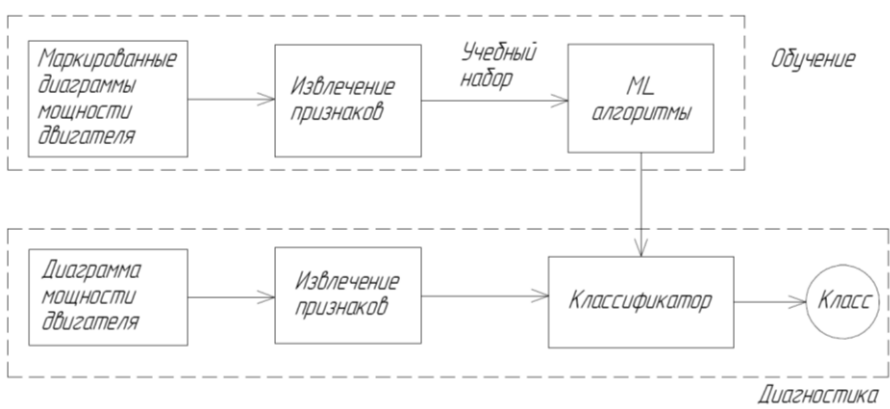


Рис. 12. Диагностическая модель

Fig. 12. Diagnostic model

На рис. 13–16 представлены матрицы распознавания, показывающие результаты классификации, диагональные элементы в этих матрицах представляют собой точно предсказанные выборки. Здесь обозначено: True Class – истинный класс; Predicted Class – предсказанный класс; TPR – доля правильно классифицированных наблюдений на истинный класс; FNR – доля неправильно классифицированных наблюдений на истинный класс. Строки с первой по шестую соответствуют рабочим состояниям штангового насоса: нормальное, негерметичность нагнетательного клапана, эффект захвата газа, недостаточное поступление жидкости, плунжер достигает верхней мертвой точки и плунжер достигает нижней мертвой точки соответственно.

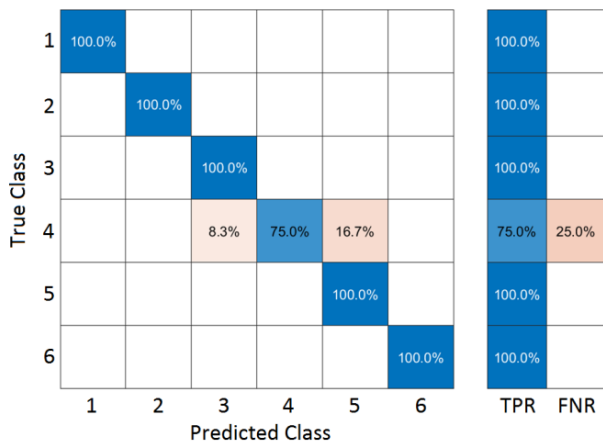


Рис. 13. Матрица распознавания для метода дерева решений

Fig. 13. Confusion matrix of Decision tree method

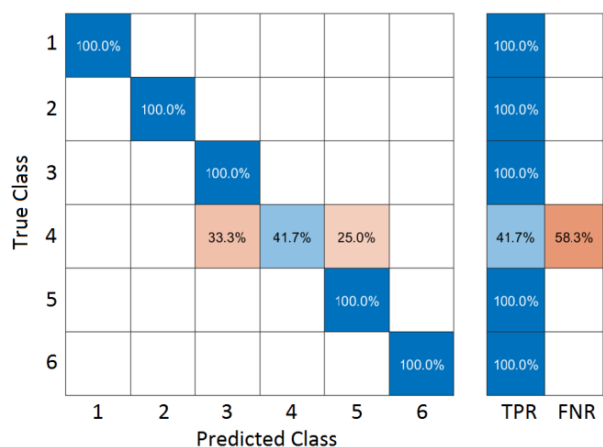


Рис. 14. Матрица распознавания для наивного байесовского классификатора

Fig. 14. Confusion matrix of Naive Bayes classifier

Как видно на рис. 13, метод дерева решений с очень высокой точностью (95,8 %) правильно клас-

сифицировал все образцы для нормального состояния, утечки нагнетательного клапана, эффект захвата газа, достижение плунжером верхней и нижней мертвой точек. Однако 16,7 % образцов с недостаточным поступлением жидкости ошибочно классифицируются как достижение плунжером верхней мертвой точки, а 8,3 % из них – как эффект захвата газа.

Наивный байесовский классификатор (рис. 14) является вторым по точности классификации, которая составляет 93,1 %, он также не смог правильно классифицировать все образцы с недостаточным поступлением жидкости, правильно классифицировав только 41,7 %.

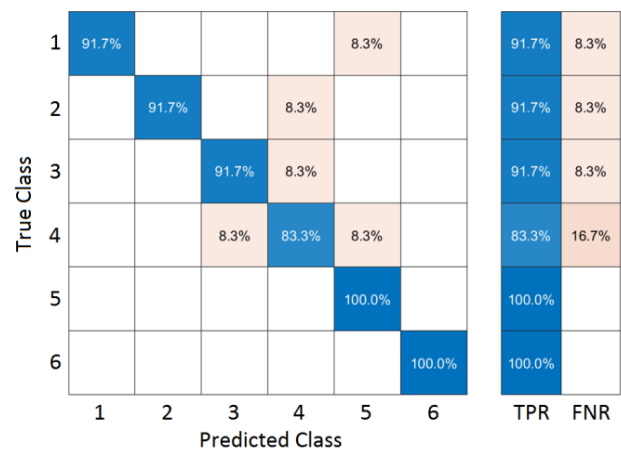


Рис. 15. Матрица распознавания для метода опорных векторов

Fig. 15. Confusion matrix of Support Vector Machine

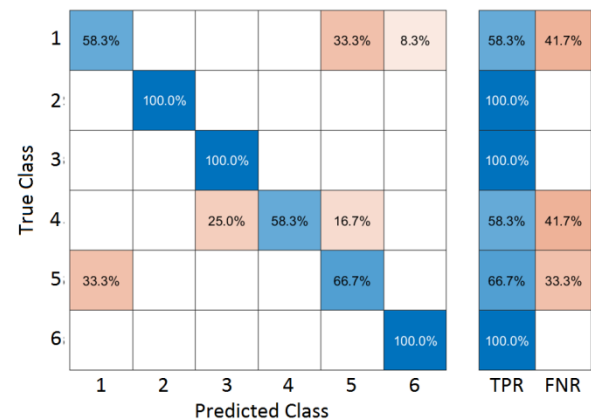


Рис. 16. Матрица распознавания для метода K-ближайших соседей

Fig. 16. Confusion matrix of K-Nearest Neighbors method

Метод опорных векторов (рис. 15) обеспечивает среднюю точность – 90,3 % – и более высокую точность правильного распознавания ситуации с

недостаточным поступлением жидкости (83,3 %), чем предыдущие методы, однако он не мог точно классифицировать все образцы для нормального состояния, негерметичность нагнетательного клапана и эффект захвата газа.

Классификатор на основе метода К-ближайших соседей (рис. 16) показал среднюю точность распознавания 80,6 %, он не смог точно классифицировать все образцы при нормальном состоянии, недостаточное поступление жидкости и достижение плунжером верхней мертвой точки.

Заключение

В этом исследовании, основанном на более ранней работе [16], для решения задачи идентификации состояний штангового насоса был использован более широкий круг методов. Помимо метода опорных векторов использованы также методы дерева решений, К-ближайших соседей и наивный

байесовский алгоритм. Это расширение набора методов предпринято для сравнения точности других классификаторов машинного обучения с результатами метода опорных векторов в контексте выявления неисправностей в штанговых насосах при использовании кривых мощности двигателя. При этом удалось классифицировать не четыре, а шесть режимов работы – добавлены состояния, когда плунжер достигает верхней и нижней мертвой точки. Использование кривых мощности двигателя позволяет построить таблицу признаков для обучения классификаторов. Результаты моделирования показывают, что предлагаемый подход позволяет весьма эффективно определять состояние штангового насоса. Метод дерева решений показал себя лучшим классификатором среди остальных, обеспечив среднюю точность 95,8 %, по сравнению с 90,3 % у метода опорных векторов.

СПИСОК ЛИТЕРАТУРЫ

1. Wilamowski B.M., Kaynak O. Oil well diagnosis by sensing terminal characteristics of the induction motor // *IEEE Transactions on Industrial Electronics*. – 2000. – Vol. 47 (5). – P. 1100–1107. DOI: 10.1109/41.873219.
2. Unveiling deep learning insights: a specialized analysis of sucker rod pump dynamographs, emphasizing visualizations and human insight / B. Martinović, M. Bijanić, D. Danilović, A. Petrović, B. Delibasić // *Mathematics*. – 2023. – Vol. 11 (23): article id. 04782. DOI: 10.3390/math11234782.
3. Fault diagnosis of sucker-rod pumping system using support vector machine / J. Feng, M. Wang, Y. Yang, F. Gao, Zh. Pan, W. Shan, Q. Liu, Q. Yang, J. Yuan // *Communications and Information Processing. Communications in Computer and Information Science*. – Heidelberg, Berlin: Springer, 2012. – Vol. 289. – P. 179–189. DOI: 10.1007/978-3-642-31968-6_22.
4. Xu P., Xu S., Yin H. Application of self-organizing competitive neural network in fault diagnosis of Sucker Rod Pump system // *Journal of Petroleum Science and Engineering*. – 2007. – Vol. 58 (1–2). – P. 43–48. DOI: 10.1016/j.petrol.2006.11.008.
5. Li K., Han Y., Wang T. A novel prediction method for down-hole working states of the beam pumping unit based on 8-directions chain codes and online sequential extreme learning machine // *Journal of Petroleum Science and Engineering*. – 2018. – Vol. 160. – P. 285–301. DOI: 10.1016/j.petrol.2017.10.052.
6. Multiple fault diagnosis of down-hole conditions of sucker-rod pumping wells based on Freeman chain code and DCA / K. Li, X. Gao, W. Yang, Y. Dai, Z. Tian // *Petroleum Science*. – 2013. – Vol. 10 (3). – P. 347–360. DOI: 10.1007/s12182-013-0283-4.
7. Разработка модели прогнозирования технического состояния нефтепромыслового энергомеханического оборудования с применением искусственной нейронной сети / И.С. Сухачев, С.В. Сидоров, В.В. Сушков, В.Р. Антропова // *Электротехнические и информационные комплексы и системы*. – 2022. – Т. 18. – № 1. – С. 88–94. DOI: 10.17122/1999-5458-2022-18-1-88-94.
8. Wei J., Gao X. Fault diagnosis of sucker rod pump based on deep-broad learning using motor data // *IEEE Access*. – 2020. – Vol. 8. – P. 222562–222571. DOI: 10.1109/ACCESS.2020.3036078.
9. Zheng B., Gao X., Li X. Fault detection for sucker rod pump based on motor power // *Control Engineering Practice*. – 2019. – Vol. 86. – P. 37–47. DOI: 10.1016/j.conengprac.2019.02.001.
10. Разработка интеллектуальной станции управления для установок штанговых глубинных насосов / М.Г. Пачин, А.Н. Яшин, А.С. Бодылев, М.И. Хакимьянов // *Известия Томского политехнического университета. Инжиниринг георесурсов*. – 2022. – Т. 333. – № 3. – С. 68–75. DOI: 10.18799/24131830/2022/3/3465.
11. Tracking the multi-well surface dynamometer card state for a sucker-rod pump by using a particle filter / A. Wang, G. Gong, R. Shen, W. Mao, H. Lu, K. Wang, J. Wang // *IET Communications*. – 2018. – Vol. 12 (16). – P. 2058–2066. DOI: 10.1049/iet-com.2018.5331.
12. Chen L., Gao X., Li X. Using the motor power and XGBoost to diagnose working states of a sucker rod pump // *Journal of Petroleum Science and Engineering*. – 2021. – Vol. 199 (6): article id. 108329. DOI: 10.1016/j.petrol.2020.108329.
13. Fault diagnosis of sucker rod pumping system using modified extreme learning machine assisted by gravitational search algorithm / X. Deng, Y. Gao, M. Yang, Z. Sun, W. Cui // *2019 IEEE Chinese Control And Decision Conference (CCDC)*. – Nanchang, China, 3–5 June 2019. DOI: 10.1109/CCDC.2019.8833177.
14. Ramez A., Abu El Ela M., El-Banbi A. Identification of downhole conditions in sucker rod pumped wells using deep neural networks and genetic algorithms (includes associated discussion) // *SPE Production & Operations*. – 2020. – Vol. 35.02. – P. 435–447. DOI: 10.2118/200494-PA.
15. Liu J., Feng J., Gao X. Fault diagnosis of rod pumping wells based on support vector machine optimized by improved chicken swarm optimization // *IEEE Access*. – 2019. – Vol. 7. – P. 171598–171608. DOI: 10.1109/ACCESS.2019.2956221.
16. Teclé S.I., Ziuzev A., Spiridonov V. Development of fault diagnosis model for sucker rod pump using motor power curve. // *2021 XVIII International Scientific Technical Conference Alternating Current Electric Drives (ACED)*. – Ekaterinburg, Russia, IEEE, 24–27 May 2021. – P. 1–6. DOI: 10.1109/ACED50605.2021.9462307.

17. Takács G. Sucker-rod pumping manual. – Tulsa, Oklahoma: PennWell Books, 2003. – 395 p.
18. Fault diagnosis for down-hole conditions of sucker rod pumping systems based on the FBH–SC method / K. Li, X.-W. Gao, H.-B. Zhou, Y. Han // *Petroleum Science*. – 2015. – Vol. 12. – P. 135–147. DOI: 10.1007/s12182-014-0006-5.
19. Quinlan J.R. C4.5: programs for machine learning. – San Francisco: Morgan Kaufmann Publ., Inc., 1993. – 302 p.
20. Ngoc-Tu N., Kwon J.-M., Lee H.-H. Fault diagnosis of induction motor using decision tree with an optimal feature selection // 2007 7th International Conference on Power Electronics. – Daegu, Korea, 27–30 November 2007. DOI: 10.1109/ICPE.2007.4692484.
21. He D., Li R., Zhu J. Plastic bearing fault diagnosis based on a two-step data mining approach // *IEEE Transactions on Industrial Electronics*. – 2013. – Vol. 60 (8). – P. 3429–3440. DOI: 10.1109/TIE.2012.2192894.
22. Jamil M.A., Khan M.A.A., Khanam S. Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults // *Vibroengineering PROCEDIA*. – 2021. – Vol. 39. – P. 36–42. DOI: 10.21595/vp.2021.22307.

Информация об авторах

Осман Хассан Ахмед, аспирант кафедры электропривода и автоматизации промышленных установок Уральского федерального университета имени первого Президента России Б.Н. Ельцина, Россия, 620002, г. Екатеринбург, ул. Мира, 19. ahmed3.othman3@gmail.com

Самуэль Исаак Текле, PhD, ассистент, Колледж инженерии и технологий, Май Нефти, Эритрея, samuel47tecle@gmail.com

Анатолий Михайлович Зюев, доктор технических наук, профессор кафедры электропривода и автоматизации промышленных установок Уральского федерального университета имени первого Президента России Б.Н. Ельцина, Россия, 620002, г. Екатеринбург, ул. Мира, 19. a.m.zyuzev@urfu.ru; <https://orcid.org/0000-0002-2233-2730>

Владимир Павлович Метельков, доктор технических наук, профессор кафедры электропривода и автоматизации промышленных установок Уральского федерального университета имени первого Президента России Б.Н. Ельцина, Россия, 620002, г. Екатеринбург, ул. Мира, 19. v.p.metelkov@urfu.ru; <https://orcid.org/0000-0001-5640-1637>

Поступила в редакцию: 07.03.2024

Поступила после рецензирования: 20.03.2024

Принята к публикации: 05.11.2024

REFERENCES

1. Wilamowski B.M., Kaynak O. Oil well diagnosis by sensing terminal characteristics of the induction motor. *IEEE Transactions on Industrial Electronics*, 2000, vol. 47, no. 5, pp. 1100–1107. DOI: 10.1109/41.873219.
2. Martinović B., Bijanić M., Danilović D., Petrović A., Delibasić B. Unveiling deep learning insights: a specialized analysis of sucker rod pump dynamographs, emphasizing visualizations and human insight. *Mathematics*, 2023, vol. 11, no. 23, article id. 04782, pp. 1–29. DOI: 10.3390/math11234782.
3. Feng J., Wang M., Yang Y., Gao F., Pan Zh., Shan W., Liu Q., Yang Q., Yuan J. Fault diagnosis of sucker-rod pumping system using support vector machine. *Communications and Information Processing. Communications in Computer and Information Science*. Berlin, Heidelberg, Springer, 2012. Vol. 289, pp. 179–189. DOI: 10.1007/978-3-642-31968-6_22.
4. Xu P., Xu S., Yin H. Application of self-organizing competitive neural network in fault diagnosis of Sucker Rod Pump system. *Journal of Petroleum Science and Engineering*, 2007, vol. 58 (1–2), pp. 43–48. DOI: 10.1016/j.petrol.2006.11.008.
5. Li K., Han Y., Wang T. A novel prediction method for down-hole working states of the beam pumping unit based on 8-directions chain codes and online sequential extreme learning machine. *Journal of Petroleum Science and Engineering*, 2018, vol. 160, pp. 285–301. DOI: 10.1016/j.petrol.2017.10.052.
6. Li K., Gao X., Yang W., Dai Y., Tian Z. Multiple fault diagnosis of down-hole conditions of sucker-rod pumping wells based on Freeman chain code and DCA. *Petroleum Science*, 2013, vol. 10 (3), pp. 347–360. DOI: 10.1007/s12182-013-0283-4.
7. Sukhachev I.S., Sidorov S.V., Sushkov V.V., Antropova V.R. Development of a model for forecasting the technical condition of oilfield power-mechanical equipment using an artificial neural network. *Electrical and data processing facilities and systems*, 2022, vol. 18, no. 1, pp. 88–94. (In Russ). DOI: 10.17122/1999-5458-2022-18-1-88-94.
8. Wei J., Gao X. Fault diagnosis of sucker rod pump based on deep-broad learning using motor data. *IEEE Access*, 2020, vol. 8, pp. 222562–222571. DOI: 10.1109/ACCESS.2020.3036078.
9. Zheng B., Gao X., Li X. Fault detection for sucker rod pump based on motor power. *Control Engineering Practice*, 2019, vol. 86, pp. 37–47. DOI: 10.1016/j.conengprac.2019.02.001.
10. Pachin M.G., Yashin A.N., Bodayev A.S., Khakimyanov M.I. Development of an intelligent control station for rod pumps. *Bulletin of the Tomsk Polytechnic University, Geo Assets Engineering*, 2022, vol. 333, no. 3. pp. 68–75. (In Russ). DOI: 10.18799/24131830/2022/3/3465.
11. Wang A., Gong G., Shen R., Mao W., Lu H., Wang K., Wang J. Tracking the multi-well surface dynamometer card state for a sucker-rod pump by using a particle filter. *IET Communications*, 2018, vol. 12 (16), pp. 2058–2066. DOI: 10.1049/iet-com.2018.5331.
12. Chen L., Gao X., Li X. Using the motor power and XGBoost to diagnose working states of a sucker rod pump. *Journal of Petroleum Science and Engineering*, 2021, vol. 199, article id. 108329. DOI: 10.1016/j.petrol.2020.108329.

13. Deng X., Gao Y., Yang M., Sun Z., Cui W. Fault diagnosis of sucker rod pumping system using modified extreme learning machine assisted by gravitational search algorithm. *2019 IEEE Chinese Control And Decision Conference (CCDC)*. Nanchang, China, 3–5 June 2019. DOI: 10.1109/CCDC.2019.8833177.
14. Ramez A., Abu El Ela M., El-Banbi A. Identification of downhole conditions in sucker rod pumped wells using deep neural networks and genetic algorithms (includes associated discussion). *SPE Production & Operations*, 2020, vol. 35.02, pp. 435–447. DOI: 10.2118/200494-PA.
15. Liu J., Feng J., Gao X. Fault diagnosis of rod pumping wells based on support vector machine optimized by improved chicken swarm optimization. *IEEE Access*, 2019, vol. 7, pp. 171598–171608. DOI: 10.1109/ACCESS.2019.2956221
16. Teclé S.I., Ziuzev A., Spiridonov V. Development of fault diagnosis model for sucker rod pump using motor power curve. *2021 XVIII International Scientific Technical Conference Alternating Current Electric Drives (ACED)*. Ekaterinburg, Russia, IEEE, May 24–27, 2021. pp. 1–6. DOI: 10.1109/ACED50605.2021.9462307.
17. Takács G. *Sucker-rod pumping manual*. Tulsa, Oklahoma, PennWell Books, 2003. 395 p.
18. Li K., Gao X.-W., Zhou H.-B., Han Y. Fault diagnosis for down-hole conditions of sucker rod pumping systems based on the FBH-SC method. *Petroleum Science*, 2015, vol. 12, pp. 135–147. DOI: 10.1007/s12182-014-0006-5.
19. Quinlan J.R. *C4.5: Programs for Machine Learning*. San Francisco, Morgan Kaufmann Publ., Inc., 1993. 302 p.
20. Ngoc-Tu N., Kwon J.-M., Lee H.-H. Fault diagnosis of induction motor using decision tree with an optimal feature selection. *2007 7th International Conference on Power Electronics*. Daegu, Korea, 27–30 November 2007. DOI: 10.1109/ICPE.2007.4692484.
21. He D., Li R., Zhu J. Plastic bearing fault diagnosis based on a two-step data mining approach. *IEEE Transactions on Industrial Electronics*, 2013, vol. 60 (8), pp. 3429–3440. DOI: 10.1109/TIE.2012.2192894.
22. Jamil M.A., Khan M.A.A., Khanam S. Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults. *Vibroengineering PROCEDIA*, 2021, vol. 39, pp. 36–42. DOI: 10.21595/vp.2021.22307.

Information about the authors

Othman H. Ahmed, Postgraduate Student, Ural Federal University named after the first President of Russia B.N. Yeltsin, 19, Mira street, Ekaterinburg, 620002, Russian Federation. ahmed3.othman3@gmail.com

Samuel I. Teclé, PhD, Assistant Professor, College of Engineering and Technology, Mai Nefhi, Eritrea, samuel47teclé@gmail.com

Anatoliy A. Zyuzev, Dr. Sc., Professor, Ural Federal University named after the first President of Russia B.N. Yeltsin, 19, Mira street, Ekaterinburg, Russian Federation. a.m.zyuzev@urfu.ru; <https://orcid.org/0000-0002-2233-2730>

Vladimir P. Metelkov, Dr. Sc., Professor, Ural Federal University named after the first President of Russia B.N. Yeltsin, 19, Mira street, Ekaterinburg, Russian Federation. v.p.metelkov@urfu.ru; <https://orcid.org/0000-0001-5640-1637>

Received: 07.03.2024

Revised: 20.03.2024

Accepted: 05.11.2024