

СОВЕРШЕНСТВОВАНИЕ ОПТИКОЭЛЕКТРОННЫХ СИСТЕМ УПРАВЛЕНИЯ НА БАЗЕ НПВО С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Кенгесбаев Салават Куанишбаевич

Докторант кафедры “Системы телерадиовещания” ТУИТ имени Мухаммада
ал-Хоразмий

salawatkenesbaev@gmail.com

Аннотация: Исследует важную роль машинного обучения (МО) в разработке методологии анализа данных в многофункциональных оптоэлектронных системах. По мере того, как эти системы становятся все более сложными, интеграция методов МО становится все более важной для получения значимого понимания из обширных наборов данных, генерируемых оптоэлектронными устройствами. Алгоритмы МО облегчают обработку в реальном времени, распознавание образов и прогнозное моделирование, повышая эффективность и точность анализа данных. Автоматизируя извлечение сложных паттернов и корреляций, мо’ позволяет исследователям обнаруживать скрытые отношения между различными оптоэлектронными параметрами. В этой статье предлагается алгоритм К-ближайших соседей (KNN) для сравнения значений и анализа данных оптоэлектронного колориметра, метод, который может эффективно решить проблему шума и способствовать повышению эффективности классификации.

Ключевые слова: МО, KNN, НПВО, DT, NB, RF, анализ данных, оптоэлектронные системы, прогнозное моделирование.

1. ВВЕДЕНИЕ

В последнее время методы оптического контроля успешно применяются для количественного и качественного анализа веществ. Их преимуществами перед другими методами физико-технического анализа являются высокая точность, чувствительность и эффективность.

Суть метода оптического анализа состоит в том, что любое вещество отражает, поглощает или излучает свет. Кроме того, в зависимости от химического состава вещества и количественного соотношения входящих в него элементов изменяются интенсивность поглощения света, угол отражения и испускания света и другие свойства взаимодействия вещества.

Оптические приборы для контроля физико-химических параметров жидкой среды можно разделить на три группы в зависимости от физических явлений, изменяющих параметры светового потока, проходящего через исследуемую среду: **фотометрические, рефрактометрические, поляриметрические** [1].

Метод поляриметрического анализа продуктов основан на измерении угла поворота плоскости поляризации света, проходящего через оптически активную среду.

Поляриметрический анализ применяется для исследования сред, содержащих оптически активные вещества. К таким веществам относятся, например, сахар, кварцевое стекло, глюкоза и др., которые способны изменять плоскость поляризации проходящего через них поляризованного света. Ограничение количества оптически активных веществ препятствует широкому применению этого метода[2].

Фотометрический метод в основном применяют для анализа жидких веществ и растворов. Этот метод сравнивает свет анализируемой жидкости со стандартным раствором или оценивает уровень поглощения света при измерении [3].

Рефрактометрический метод, основанный на определении показателя преломления исследуемой среды с использованием эффектов нарушенного полного внутреннего отражения (НПВО) и многократного нарушенного полного внутреннего отражения (МНПВО), характеризуется высокой точностью, технической простотой и доступностью [4].

С точки зрения обработки данных текущий оптико-электронные измерительные системы требуют много времени для анализа физико-химических параметров жидкой среды, а ограниченная автоматизация замедляет обработку данных. То есть традиционные системы не обладают скоростью обработки и пропускной способностью данных, необходимой для анализа в реальном времени. Они часто не могут адаптироваться к изменяющимся условиям в режиме реального времени, что приводит к задержкам и недостаточному времени отклика для быстрого анализа и принятия решений.

В настоящее время оптоэлектронные измерительные системы нуждаются в передовых решениях для обработки данных, точности и анализа физико-химических параметров жидких сред в реальном времени. Интеграция машинного обучения с оптоэлектронными измерительными системами направлена на устранение ограничений современных систем. Применяя алгоритмы машинного обучения, измерительные системы решают проблемы, связанные с обработкой данных, точностью и анализом в реальном времени, что приводит к повышению эффективности и производительности.

Целью анализа физико-химических параметров жидких сред является применение алгоритма KNN для прогнозирования параметров пробы с помощью рефрактометрического метода.

II. МЕТОДЫ

В настоящее время оптоэлектронные системы стали незаменимыми инструментами в различных областях. Эти системы используют взаимодействие между светом и материей для управления и измерения широкого спектра физических и химических свойств. Среди многочисленных оптических явлений нарушенного полного внутреннего отражения (НПВО) выделяется как очень универсальный и мощный механизм исследования

свойств материалов. Использование алгоритмов машинного обучения дает возможность совершенствовать оптоэлектронные системы.

А. Метод и элемент нарушенного полного внутреннего отражения (НПВО)

В последнее время оптико-электронные методы контроля успешно применяются для количественного и качественного анализа веществ. Высокая точность и чувствительность, эффективность являются их преимуществами перед другими методами физико-технического анализа.

Нисколько не уступая другим физико-химическим методам в точности и удобстве определения оптических параметров метод нарушенного полного внутреннего отражения (НПВО) превосходит их в экспрессности, обеспечивая в то же время неразрушающий контроль.

Приставки НПВО (рис. 1.) выпускаются по двум типовым схемам. Схема а) наиболее проста, так как в ней использованы только плоские зеркала и простой призматический элемент НПВО.

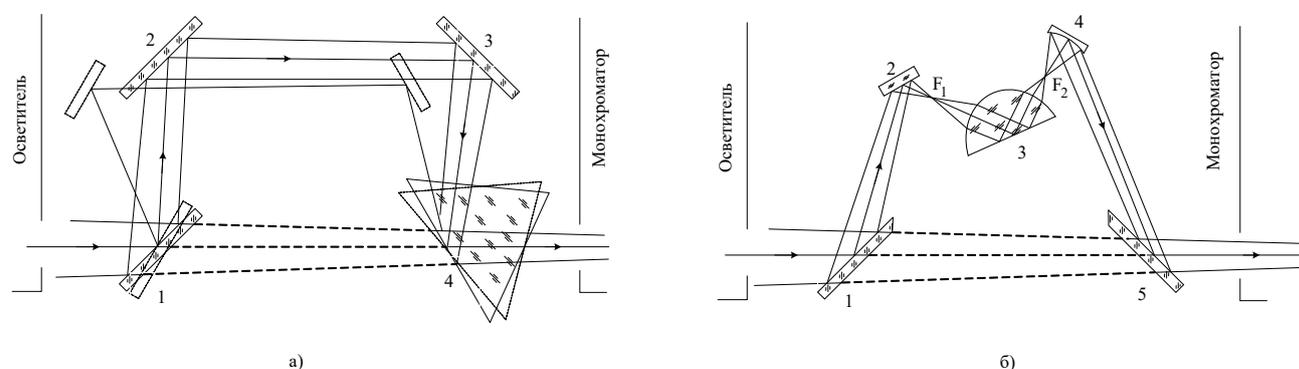


Рис.1. Приставки НПВО:

- а) – призматическая: 1–3 – плоские зеркала; 4 – элемент НПВО;
- б) – полуцилиндрическая: 1, 5 – плоские зеркала; 2, 4 - торoidalные зеркала; 3 – элемент НПВО

Схема б) с полуцилиндрическим элементом НПВО более совершенна, поскольку в ней обеспечивается близкий к параллельному пучок света внутри элемента, но несколько более сложна по конструкции[5].

В основе метода НПВО (рис. 2) лежит явление проникновения световой волны в оптически менее плотную среду n_2 , когда световой поток $\Phi_{0\lambda_1}$ распространяется из оптически более плотной среды n_1 в менее плотную n_2 под углом Θ .

Физический смысл заключается в следующем. Используются два световых потока с разными длинами волн, например, λ_1 (измерительный) и λ_2 (опорный), соответствующим максимальному и минимальному поглощению жидкость.

Световой поток $\Phi_{0\lambda}$ распространяется из оптически плотной среды с показателем преломления n_1 , в менее плотную с n_2 под углом Θ , превышающим предельный (критический) угол полного внутреннего отражения Θ_k . Световой поток $\Phi_{0\lambda_1}$ частично проникает в среду n_2 ; на глубину порядка длины волны падающего излучения [6]:

$$d_p = \frac{\lambda_1}{2\pi(\sin^2 \Theta - n_{21})^{1/2}} \quad (1)$$

где λ_1 – длина волны излучения в оптически плотной среде с показателем преломления n_1 ; $n_{21} = n_2/n_1$ – относительный показатель преломления. Поскольку менее плотная среда с комплексным показателем преломления $\tilde{n} = n_2 - j\chi_2$: обладает поглощением, отражение не будет полным, т. е. условия для этого нарушаются и коэффициент отражения ($R = \Phi_{\lambda_1} / \Phi_{0\lambda_1}$) станет меньше 1. Степень ослабления R пропорциональна показателю поглощения. Таким образом, чем выше поглощение, тем сильнее нарушается отражение. Это называется эффект НПВО.

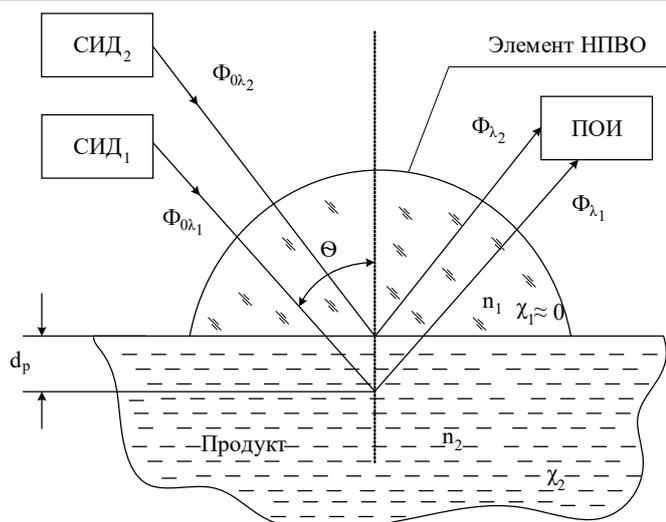


Рис.2. Физические основы метода НПВО

В ИК-области спектра для достижения условия $n_1 > n_2$ используются измерительные элементы НПВО из высокопреломляющих оптических материалов, прозрачных в соответствующем диапазоне[7].

Основным законом поглощения света, основанным на многих аналитических методах, является закон Бугера-Ламберта-Бера [8]. Он определяет ослабление пучка света при его прохождении через поглощающее вещество. Если на слой вещества (в нашем случае на контролируемую жидкую среду) толщиной d падает параллельный пучок света с интенсивностью Φ_{01} , то на выходе из слоя интенсивность пучка Φ будет равна

$$\Phi = \Phi_{01} e^{-kCd} \quad (2)$$

где k – показатель поглощения, зависящий от длины волны; C – концентрация анализируемого вещества.

Φ_{01} - световой поток до прохождения через слой контролируемой жидкости;

Φ - световой поток после прохождения через слой контролируемой жидкости.

В нашем случае отраженный световой поток равен Φ_{λ_1} и Φ_{λ_2} :

$$\begin{cases} \Phi_{\lambda_1} = \Phi_{0\lambda_1} e^{-(k_2Cd+k_2Cd)} = \Phi_{0\lambda_1} e^{-2k_2Cd} \\ \Phi_{\lambda_2} = \Phi_{0\lambda_2} \end{cases} \quad (3)$$

где $\Phi_{0\lambda_1}$ и $\Phi_{0\lambda_2}$ - световые потоки до прохождения через слой контролируемой жидкости; Φ_{λ_1} и Φ_{λ_2} - световые потоки после прохождения через слой контролируемой жидкости; k_2 - коэффициент поглощения в зависимости от длины волны.

Если интенсивность, соответствующая световому потоку с длинами волн λ_1 и λ_2 до прохождения через образец, одинакова, т.е.

$$\Phi_{0\lambda_1} = \Phi_{0\lambda_2} \quad (4)$$

тогда отношение интенсивностей соответствующих световых потоков с длинами волн λ_1 и λ_2 пропорционально концентрации отмеченной жидкости C .

$$\frac{\Phi_{\lambda_2}}{\Phi_{\lambda_1}} = \frac{\Phi_{0\lambda_2}}{\Phi_{0\lambda_1} e^{-2k_2Cd}} = e^{-2k_2Cd} \quad (5)$$

С помощью оптоэлектронных преобразователей можно определять те или иные оптические свойства образца по его размерам путем преобразования световых потоков в соответствующие электрические сигналы [9].

Б. Выбор модели машинного обучения

Это отрасль искусственного интеллекта, которая предполагает, что компьютеры обучаются и развивают собственный опыт без конкретных инструкций. После обучения машинному обучению компьютерный алгоритм может применять полученные знания для решения аналогичных задач. В данном исследовании мы рассматриваем несколько методов машинного обучения для оценки физико-химических параметров жидких сред на основе баз данных.

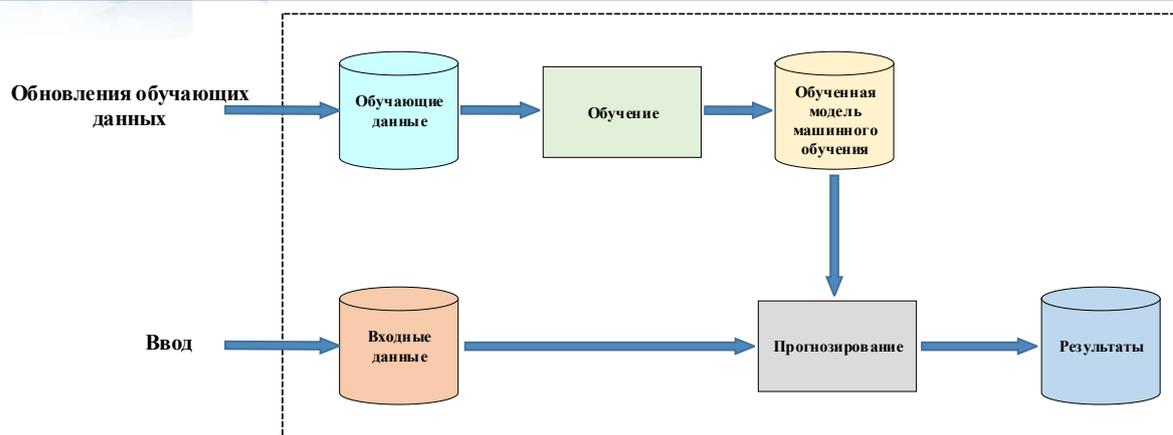


Рис. 3. Общий рабочий процесс алгоритма МО

K-ближайших соседей (K-nearest neighbors, KNN)

KNN непараметрический алгоритм, используемый для регрессии или классификации. Это простой алгоритм, который легко реализовать, а построение модели не требует больших затрат. Он использует набор данных, в котором точки данных сгруппированы в разные классы, и алгоритм стремится классифицировать данную выборочную точку данных как проблему классификации. Классификационным показателем KNN является класс объекта, который определяется суммой голосов K ближайших соседей[10].

Деревья решений (Decision trees, DT)

Дерево решений — это алгоритм машинного обучения, который используют для классификации, регрессии и выявления аномалий. Деревья решений часто применяют для анализа больших данных. Предсказания модели довольно точные и их легко визуализировать. Алгоритм сам генерирует правила в процессе обучения. Правила — это обобщения множества отдельных наблюдений, описывающих предметную область. Дерево решений анализирует выборку входящих данных, а затем группирует их, чтобы объекты одного класса оказались вместе[11].

Случайный лес (Random forest, RF)

RF алгоритм обучения с учителем. Его можно применять как для классификации, так и для регрессии. Также это наиболее гибкий и простой в

использовании алгоритм. RF создает деревья решений для случайно выбранных семплов данных, получает прогноз от каждого дерева и выбирает наилучшее решение посредством голосования. Он также предоставляет довольно эффективный критерий важности показателей (признаков)[12].

Наивный Байес (Naive Bayes, NB)

Наивный байесовский классификатор в машинном обучении — это контролируемое машинное обучение для задач классификации. NB используется для решения задач классификации. Он прогнозирует на основе вероятности объекта. NB основан на теореме Байеса и в основном используется для классификации текста. NB — это алгоритм вероятностной классификации, который легко реализовать и быстро обучать.

Поскольку наивный байесовский классификатор основан на теореме Байеса, он известен как вероятностный классификатор. Он прогнозирует на основе вероятности элемента[13].

Оценка данных физико-химических параметров жидкой среды на основе данных, собранных на основе различных параметров, визуализируется для точности и времени обучения, а результаты суммируются в таблице 1 . Точность каждого из используемых алгоритмов, таких как KNN, DT и т. д., показана в этих результатах.

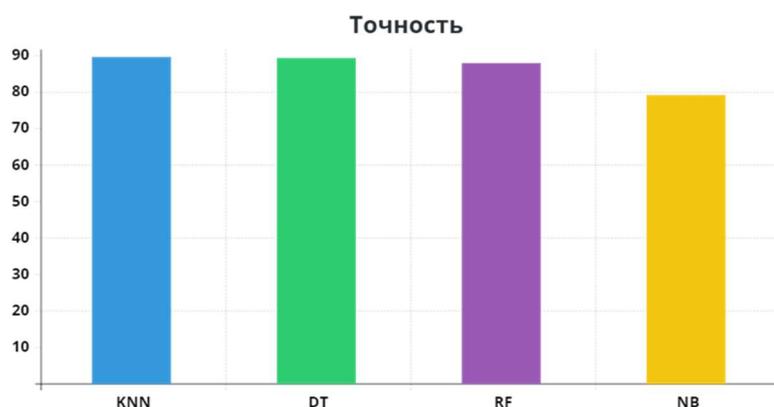


Рис. 4. Точность алгоритмов машинного обучения

Таблица 1. Результаты внедрения

Алгоритмы	Точность
KNN	88.25

Деревья решений (DT)	88.23
Случайный лес (RF)	86.82
Наивный Байес (NB)	78.06

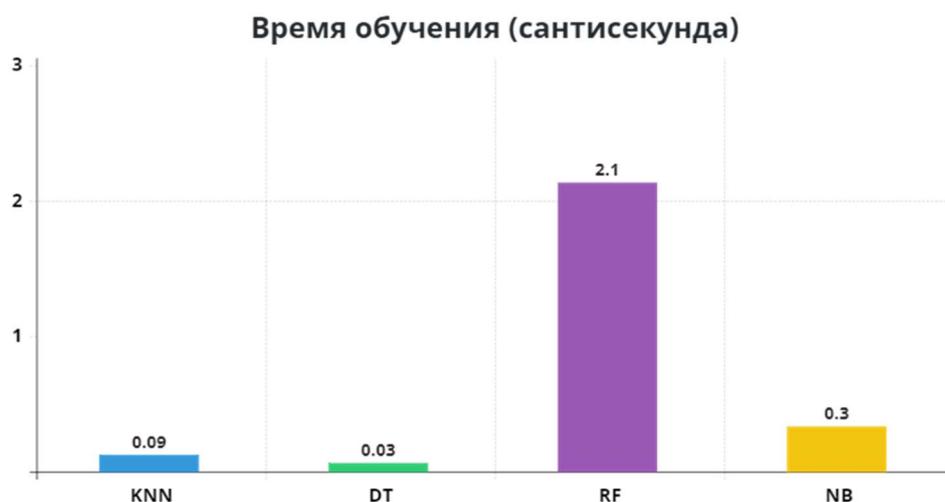


Рис. 5. Время обучения алгоритмов машинного обучения

Используя приведенные выше результаты, KNN показывает хорошую производительность по сравнению с другими алгоритмами. Но в целом производительность алгоритмов зависит от предметной области и набора данных, в которых они применяются. Алгоритмы машинного обучения могут работать лучше других при определенных условиях.

III. ПРЕДЛАГАЕМАЯ СИСТЕМА

Предлагаемая система предусматривает использование машинного обучения для значительного повышения точности и эффективности оптоэлектронных систем управления на основе НПВО. Цель состоит в том, чтобы найти алгоритм машинного обучения для классификации свойств жидкой среды на основе набора данных. KNN используется для прогнозирования параметров посредством обучения модели. Этот метод используется для обработки данных в реальном времени, а также для повышения эффективности и принятия решений.

Предлагаемая система показывает следующий полный порядок разрешения на рисунке 6 :

- Сбор данных: извлечение и импорт данных.
- Предварительная обработка данных: очистка данных и извлечение/отбор признаков.
- Обучение машинному обучению: KNN.
- Оценка модели: Тестирование.
- Выходные данные: категорическое прогнозирование.



Рис. 6. Предлагаемая архитектура системы

Методология рекомендуемых подходов. KNN — это алгоритм классификации, основанный на сходстве признаков. Он изучает данные, анализирует их удаленность и сходство и использует значения K для разделения их на кластеры. Расстояние можно определить по-разному. Для этого исследования мы будем использовать Евклидову меру расстояния. Процедуры, представленные на рисунке 7, помогут вам понять, как это работает.

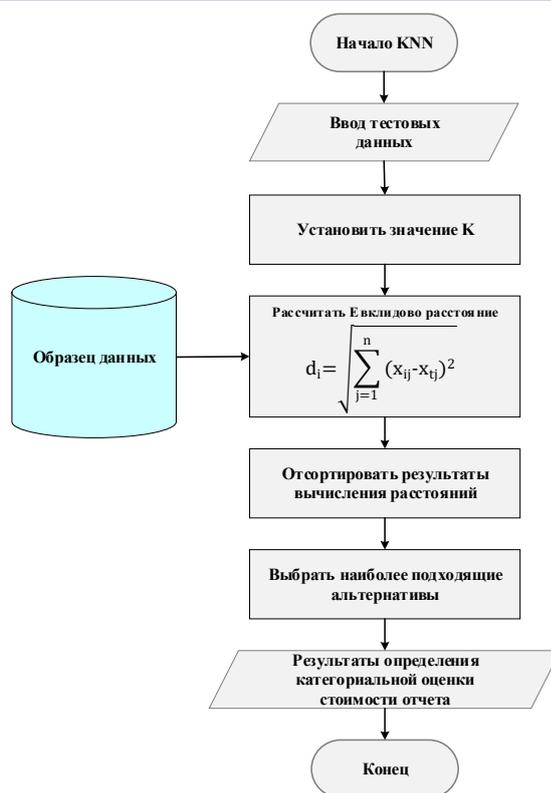


Рис. 7. Блок-схема KNN

Предлагается устройство мониторинга жидкой среды на основе использования эффекта НПВО. Блок-схема устройства показана на рисунке 8.

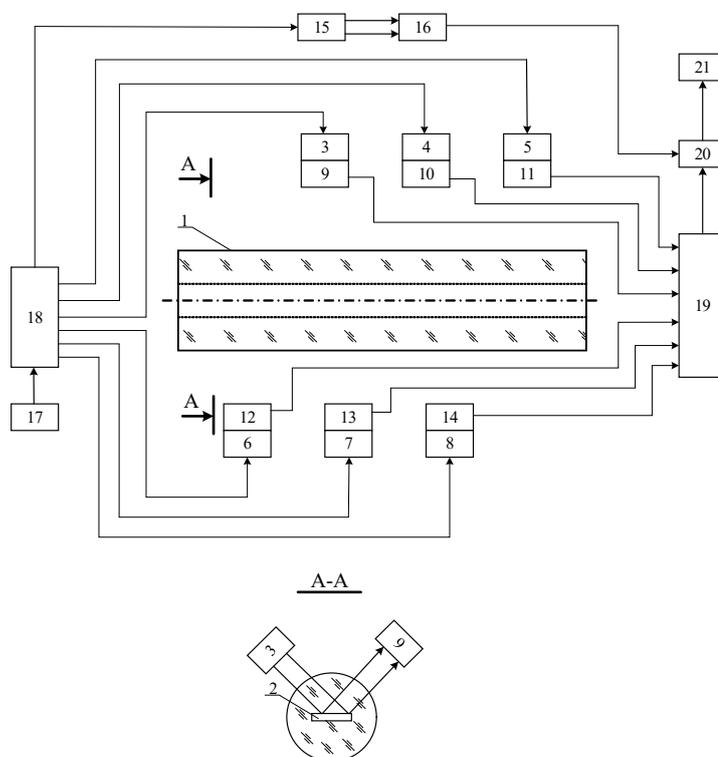


Рис. 8. Блок-схема оптоэлектронного устройства для мониторинга жидких сред

Устройство содержит цилиндр 1 с полостью в виде призмы 2, проходящей через его центр, задающий генератор (источник импульсного питания) 17, триггер 18 со счетным входом (счетчик), светоизлучающие диоды (СИД) 3–8 и компенсационный СИД 15, измерительные 9–14 и компенсационный 16 приемники оптического излучения (ПОИ), сумматор 19, блок обработки фотоэлектрического сигнала 20 и регистрирующий прибор или ЭВМ 21.

Система работает следующим образом. При включении задающего генератора 17 вырабатываемые им прямоугольные импульсы с частотой 8...10 кГц подаются на вход триггера 18. Разделенные импульсы поступают на СИД 3–8 измерительных каналов и СИД 15 опорного канала. Поток излучения светоизлучающих диодов фокусируется на измерительных поверхностях и подается на фотоприемники 9–14. Далее сигналы с фотоприемников суммируются в сумматоре 19. Излучение компенсационного светоизлучающего диода через световод попадает на компенсационный фотоприемник 16. Затем сигнал с компенсационного фотоприемника поступает в блок обработки фотоэлектрического сигнала 20, где измеряется отношение компенсационного и измерительного сигналов. Сигнал отношения, несущий информацию о содержании нефти в среде, подается на регистрирующий прибор или ЭВМ 21, по показанию которого судят о оценены оптические параметры жидкой среды.

Приведены архитектура предлагаемой системы (рис.6), блок-схема KNN и блок-схемы и рабочий процесс оптоэлектронного устройства для мониторинга жидких сред.

В целом данные, полученные от оптоэлектронного устройства контроля жидких сред, считаются данными прямых испытаний. На основе этих данных тестовые данные прогнозируются с помощью алгоритма машинного обучения

KNN, а категориальный показатель определяется в соответствии со значением отчета.

IV. ЗАКЛЮЧЕНИЕ

Интеграция машинного обучения в оптоэлектронные системы управления на основе НПВО открывает новые возможности для расширенной автоматизации и интеллектуального управления в различных приложениях. Достижения в области машинного обучения позволяют быстрее обрабатывать данные датчиков, облегчают принятие решений и контроль в режиме реального времени, а также автоматизируют такие задачи, как извлечение и классификация функций, упрощение процесса разработки и снижение затрат.

В целом, предлагаемый метод интеграция машинного обучения в анализ физико-химических данных жидких сред улучшает производительность системы, что позволяет более эффективно использовать оптические данные для различных научных, промышленных и технологических целей.

Будущие работы включают многотомное измерение параметров жидкости и повышение точности измерений за счет сбора электрических сигналов от n фотоприемников и интеграция машинного обучения.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА:

- [1]. Рахимов Н.Р. Исследование фотоэлектрических свойств АФН-пленок теллурида кадмия с серебром и разработка оптоэлектронных приборов на их основе. Дисс. канд. техн. наук, Ташкент, 2001, 140 с.
- [2]. Абдуллаев А.А. и др. Различие нефти по спектрам отражения. Изв. вузов, Нефть и газ, 1979 г., № 3, с. 67–70.
- [3]. Госьков П.И. Оптоэлектронные развертывающие полупроводниковые преобразования в измерительном языке. Томск. университет, 1978 г., 149–151 с.
- [4]. Лейкин М.В., Молочников Б.И. и др. Отражательная рефрактометрия. Л.: Машиностроение, 1983 г.

- [5]. Иоффе Б.В. Рефрактометрические методы в химии. Л.: Химия, 1983 г.
- [6]. Патент РУз IAP 20030713 от 17.09.2004 г. Устройство для определения содержания эмульсионной воды в нефти и нефтепродуктах / Ш.М. Сайдахмедов; Р.Ж. Тожиев; Н.Р. Рахимов; А.Х. Хайдаров // Б.И. 2005. № 8. – С. 141.
- [7]. Рахимов, Н.Р. Оптоэлектронные системы на основе эффекта НПВО для контроля технологических параметров нефти и нефтепродуктов / Н.Р. Рахимов, Л.Ф. Парфирьев // Изв. вузов. Приборостроение, 2005. - № 10. – С. 41 – 45
- [8]. Ушаков, О.К. Автоматический колориметр на основе оптоэлектронных преобразователей/ О.К. Ушаков, Е.Ю. Кутенкова, А.Н. Серьезнов / Интерэкспо ГЕО – Сибирь-2012. VIII Междунар. науч. конгр., 10 – 20 апреля 2012 г. Новосибирск : Междунар. науч. конф. «Специализированное приборостроение, метрология, теплофизика, микротехника, нанотехнологии»: сб. материалов в 2 т. Т.1. – Новосибирск : СГГА, 2012. – С. 146 - 151.
- [9]. Н.Р.Рахимов, Б.Н.Рахимов оптоэлектронные преобразователи в нефтегазовой отрасли: монография.2018.-104-110 с.
- [10]. Zhang, S., Zhang, S., Wang, B. and Habetler, T.G., 2019. Machine Learning and Deep Learning Algorithms for Bearing Fault Diagnostics-A Comprehensive Review. arXiv preprint arXiv:1901.08247.
- [11]. Shalev-Shwartz, S. and Ben-David, S., 2014. Understanding machine learning: From theory to algorithms. Cambridge university press.
- [12]. Kotsiantis, S.B., Zaharakis, I., and Pintelas, P., 2007. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, pp.3-24.
- [13]. I. Rish, J. Hellerstein, and T. Jayram. An analysis of data characteristics that affect naive Bayes performance. Technical Report RC21993, IBM T.J. Watson Research Center, 2001.