

**М. С. Загорникова**  
(ГГУ имени Ф. Скорины, Гомель)  
Науч. рук. **С. А. Лукашевич**, ст. преподаватель

## **ИСПОЛЬЗОВАНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ФИЗИЧЕСКИХ ДАННЫХ**

Современная физика сталкивается с необходимостью обработки огромных объемов данных (временные ряды, спектры, изображения), которые сложно анализировать традиционными методами. Машинное обучение (МО) предлагает инструменты для автоматизации анализа, поиска закономерностей и прогнозирования, что особенно важно для специальности «Компьютерная физика», объединяющей физику, математику и информационные технологии (ИТ).

Целью работы является систематизация подходов к применению МО для анализа физических данных в рамках учебной программы «Компьютерная физика».

Основные задачи, поставленные в работе:

- изучить теоретические основы МО: алгоритмы, метрики, этапы работы;
- проанализировать особенности физических данных (типы, предобработка);
- рассмотреть методы, библиотеки и инструменты для анализа.
- показать интеграцию МО в дисциплины (параллельные вычисления, моделирование).

Ключевая идея заключается в том, что машинное обучение перестает быть просто вспомогательным инструментом – оно становится частью методологии физических исследований.

Автоматизация анализа данных, построение точных прогностических моделей и даже открытие новых физических закономерностей – все это теперь возможно благодаря интеграции МО в научные процессы.

Машинное обучение – это современный подход к анализу данных, который особенно полезен при работе с большими объемами информации в физике. В отличие от традиционных методов, где все правила задаются вручную, здесь компьютер сам находит закономерности в данных.

Основные направления машинного обучения можно разделить на три типа:

- обучение с учителем – когда у нас есть готовые примеры с правильными ответами. Например, можно научить компьютер распознавать разные типы частиц по данным с детекторов;
- обучение без учителя – когда мы ищем скрытые закономерности в данных без готовых ответов. Это полезно, например, для группировки звезд по их характеристикам;
- обучение с подкреплением – когда система учится методом проб и ошибок. Так можно оптимизировать параметры сложных экспериментов.

Среди популярных методов стоит отметить:

- линейную и логистическую регрессию для предсказаний [1];
- метод опорных векторов для классификации данных;
- нейронные сети для сложных задач, таких как анализ изображений.

Сравнительный график линейной и логистической регрессии показан на рисунке 1.

Принцип работы линейной регрессии методом опорных векторов показан на рисунке 2. На рисунке 3 проиллюстрирован пример простой нейронной сети.

Особенно важно правильно оценивать качество работы алгоритмов. Для этого используют различные показатели точности и проверяют, как хорошо модель работает на новых данных.

Главное преимущество машинного обучения в физике – это возможность автоматически обрабатывать большие объемы экспериментальных данных и находить в них скрытые закономерности, которые трудно заметить стандартными методами.

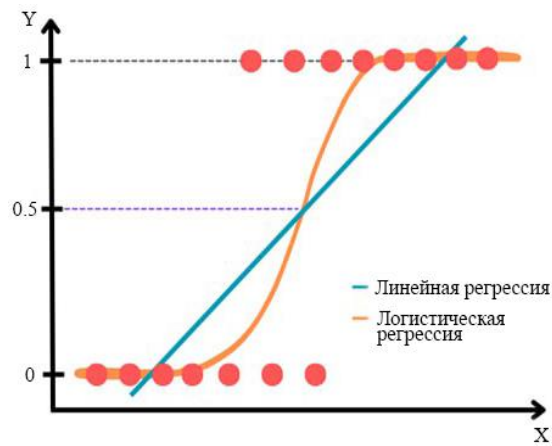


Рисунок 1 – Сравнительный график линейной и логистической регрессии

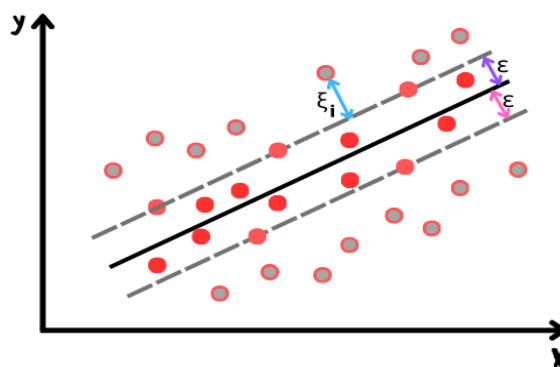


Рисунок 2 – Принцип работы линейной регрессии методом опорных векторов

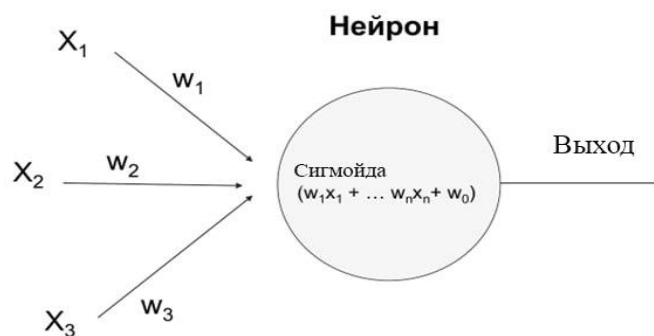


Рисунок 3 – Пример простой нейронной сети

Физические эксперименты и наблюдения генерируют огромные объемы сложных данных — от показаний датчиков до астрономических снимков. Эти данные трудно анализировать традиционными методами из-за их масштабов и особенностей. Машинное обучение предлагает новые подходы к обработке такой информации.

Перед тем как применять алгоритмы, данные необходимо тщательно подготовить. Люди очищают их от случайных ошибок и шумов, приводят к единому формату, выделяют наиболее значимые параметры. Это кропотливая работа, но без нее невозможно получить достоверные результаты [2].

Основная сложность заключается в том, что физические данные часто противоречивы и неочевидны. Приборы могут давать погрешности, а некоторые явления трудно измерить

напрямую. Кроме того, важно не просто получить ответ от компьютера, но и понять его физический смысл.

Сегодня машинное обучение помогает решать множество практических задач в физике. Алгоритмы предсказывают солнечную активность, анализируют поведение частиц в ускорителях, расшифровывают свет далеких звезд. Но самое ценное – это совместная работа программистов и физиков, когда технические решения помогают глубже понять законы природы.

Для обработки физических данных применяются специализированные вычислительные методы и программные инструменты. Python с научными библиотеками представляет собой основной рабочий инструментарий благодаря своей гибкости и производительности.

NumPy обеспечивает эффективную работу с многомерными массивами, Pandas предоставляет удобные структуры данных для табличной информации. SciPy содержит набор алгоритмов для научных вычислений, а Matplotlib позволяет качественно визуализировать результаты [3].

В области машинного обучения Scikit-learn реализует классические алгоритмы, включая методы классификации, регрессии и кластеризации. TensorFlow и PyTorch используются для построения и обучения нейронных сетей различной архитектуры.

Критически важным этапом является валидация моделей. Применяются методы кросс-валидации, анализ обучающих кривых, подбор гиперпараметров с использованием сеточного поиска и байесовской оптимизации. Особое внимание уделяется интерпретируемости результатов и их физической достоверности.

Эффективная визуализация играет ключевую роль в анализе и интерпретации физических данных. Для временных рядов применяются линейные графики с выделением трендов и аномалий. Диаграммы рассеяния позволяют выявлять корреляции между параметрами.

Тепловые карты используются для представления матричных данных, таких как корреляционные матрицы или пространственные распределения. 3D-визуализация применяется для анализа объемных данных и сложных структур.

При создании визуализаций соблюдается баланс между информативностью и наглядностью. Используются научно обоснованные цветовые палитры, четкая разметка осей и подписей. Современные инструменты позволяют создавать интерактивные визуализации для более глубокого анализа.

Машинное обучение открывает новые возможности для анализа физических данных, позволяя решать задачи, которые раньше казались слишком сложными или трудоемкими. В ходе исследования мы убедились, что современные алгоритмы могут не только ускорять обработку экспериментальных результатов, но и помогать обнаруживать скрытые закономерности.

Важно понимать, что применение этих методов требует глубокого понимания как физической сути задач, так и принципов работы алгоритмов. Успех зависит от грамотной подготовки данных, правильного выбора модели и тщательной проверки результатов. Особую ценность представляет сочетание вычислительных возможностей машинного обучения с физической интуицией исследователя.

Перспективы развития этого направления связаны с созданием более специализированных инструментов, учитывающих особенности физических данных. Также важно работать над улучшением интерпретируемости результатов, чтобы компьютерные выводы не просто давали точные предсказания, но и помогали глубже понимать физические процессы.

Практическая ценность работы подтверждается примерами успешного применения рассмотренных методов в реальных исследованиях – от анализа данных с ускорителей частиц до обработки астрономических наблюдений. Это доказывает, что машинное обучение становится неотъемлемой частью современной физики, расширяя возможности научного познания.

## Литература

1. Машинное обучение: Линейная регрессия. Теория и реализация. С нуля. На чистом Python. [Электронный ресурс]. – Режим доступа: [https://habr.com/ru/articles/ 863194/](https://habr.com/ru/articles/863194/). – Дата доступа: 17.03.2024.
2. Как работать с временными рядами. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/550978/>. – Дата доступа: 17.03.2024.
3. Фоминых, Е. И. Инструментальное программное обеспечение : учебное пособие / Е. И. Фоминых, Т. Е. Фоминых. – Минск : РИПО, 2022. – 240 с.