Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский университет
«Высшая школа экономики»

На правах рукописи

Чертенков Владислав Игоревич

Исследование универсальности моделей статистической механики методами машинного обучения

РЕЗЮМЕ ДИССЕРТАЦИИ

на соискание ученой степени кандидата наук по прикладной математике

Научный руководитель:

доктор физико-математических наук

профессор Щур Лев Николаевич

Москва – 2024

Актуальность исследования

Внедрение методов машинного обучения (МО) для исследования фазовых переходов является перспективным направлением. Алгоритмы МО позволяют решать задачи с большим объемом данных. Нейронные сети (НС) широко применяются в задачах анализа естественного языка, визуального распознавания объектов, прогнозирования временных рядов и инженерных приложениях. Для решения проблем в различных доменных областях, алгоритмам не требуется априорное понимание природы данных, а только достаточное количество примеров для обучения. Например, для понимания естественного языка, нейронной сети не нужно знать принципы морфологии и семантики языка.

Анализ количества научных публикаций с 2015 по 2023 год, по данным Google Scholar, показывает рост интереса к инструментам МО в физике конденсированного состояния. Использование машинного обучения для исследования фазовых переходов позволит объединить классические методы статистической физики и передовые подходы к моделированию сложных процессов. Междисциплинарный подход позволит обогатить текущую теоретическую базу новыми методами для более глубокого понимания сложной динамики фазовых переходов и интерпретации физических систем.

Помимо теоретического значения, практическая ценность заключается в появлении новых инструментов, програмных комплексов и методов для работы с фазовыми переходами в реальных системах. А распространение методов МО откроет возможности для создания приложений в материаловедении, физике высоких энергий и других смежных областях.

Степень разработанности темы

В работе [1] предложен подход к анализу фазовых переходов методом обучения с учителем. Решая задачу классификации фаз для модели Изинга на квадратной решетке, численно извлечена критическая температура $T_c^* = 2.266(2)$ и критический показатель $\nu = 1.0(2)$ через коллапс выходных данных НС. Переносом знания получена оценка $T_c^* = 3.65(1)$ и $\nu = 1.0(3)$ для модели Изинга с треугольной топологией решетки [2]. Решая задачу регрессии для предсказания температуры спиновых конфигураций [3] осуществляется перенос знания о фазовом переходе с модели Изинга на q-компонентную модель Поттса с точностью T_c^* до 3 знака после запятой для $q \in [2; 10]$.

Для генерации новых примеров на решетке с сохранением распределения статистик термодинамических показателей [4], используют генеративно-состязательные сети (GAN). Решением задачи понижения размерности обучения без учителя [5], извлекается критическая температура $T_c^* = 2.266(4)$ в модели Изинга. Методами МО исследуют [6—9] фазовый переход в задачах перколяции и БКТ-переход в XY- и q-компонентной часовой модели.

При сравнении точности методов извлечения величин W, вводится понятие относительной погрешности метода ϵ, % (формула 1). Например, если численная оценка критической температуры $\hat{T}_c = 2.258(5)$ для точного решения $T_c = 2.269$, то $\epsilon = 100 \cdot \max(|2.269 - 2.258|, 0.005)/2.269 = 100 \cdot \max(0.011, 0.005)/2.269 = 0.5\%$.

$$\epsilon = 100 \cdot \max\left(|W - \hat{W}|, W_e\right)/W. \tag{1}$$

Численное извлечение критической температуры T_c^* с помощью методов МО неоднократно подтверждается в разных статьях. Относительная погрешность T_c^* (далее погрешность, см. выражение 1) в большинстве работ составляет менее 1%. Критический показатель корреляционной длины ν извлекается менее постоянно и с большей погрешностью. Численные оценки ν в оригинальной работе [1] имеют относительную погрешность 20% для квадратной и 30% для треугольной решеток модели Изинга. Многие работы используют метод коллапса кривых, для которого численное решение ν лежит в широком диапазоне. В работе [9] применяется другой метод для извлечения критических показателей, однако независимого подтверждения полученным результатам нет.

Возникают вопросы, с какой точностью НС могут извлекать критические свойства решеточных спиновых моделей; какие факторы влияют на точность извлекаемых величин; возможно ли создать метод, которой будет работать для моделей в разных классах универсальности.

Цель

Разработать метод для анализа фазовых переходов в решеточных спиновых моделях методами машинного обучения с учителем.

Задачи исследования

- Разработка метода анализа фазовых переходов в решеточных спиновых моделях методами машинного обучения с учителем с использованием нейронных сетей.
- Реализация программного комплекса для исследования фазовых переходов в решеточных спиновых моделях с помощью классических методов и нейронных сетей.
- Применение метода для исследования фазовых переходов в решеточных спиновых моделях классов универсальности Изинга и четырехкомпонентного Поттса.
- Измерение точности разработанного метода и исследование факторов, влияющих на точность.
- Разработка метода переноса знания о фазовом переходе в решеточных спиновых моделях.

Научная новизна

1. Предложен метод исследования фазовых переходов в решеточных спиновых моделях на основе масштабирования функции вариации выхода нейронной

сети, обученной решать задачу бинарной классификации с учителем. По выходным данным нейронной сети систематически извлекается оценка критической температуры и критический показатель корреляционной длины с высокой точностью.

2. Предложенным методом извлечены критические показатели корреляционной длины и критическая температура для моделей Изинга, Бакстера-Ву, Поттса (q=4). Для последних двух моделей критические показатели извлечены впервые методами МО.

3. Впервые исследовано влияние архитектур нейронных сетей, гиперпараметров обучения и способов предобработки входных данных на точность извлекаемых критических показателей решеточных спиновых моделей.

4. Предложен новый метод предобработки входных данных при перекрестном обучении нейронной сети с целью переноса знания о фазовом переходе между классами универсальности Изинга и четырехкомпонентной модели Поттса.

Методология и методы

Для исследования решеточных спиновых моделей классическими подходами применялись методы Монте-Карло и конечно-мерный анализ. В разработанном методе моделирования применяются методы глубокого машинного обучения, алгоритмы компьютерного зрения, нейронные сети, обучение с учителем и методы оптимизации. Общими методами исследования являются статистический анализ и численные методы аппроксимации.

Краткое содержание работы

В **первой главе** описаны исследуемые двумерные решеточные спиновые модели: Изинга [10] на квадратной решетке, модель Бакстера-Ву на треугольной решетке и четырехкомпонентная модель Поттса на квадратной решетке. В моделях наблюдается фазовый переход второго рода между ферромагнитной (упорядоченной) и парамагнитной (неупорядоченной) фазами в критической точке T_c. Существуют точные аналитические решения: модель Изинга решена Л.Онсагером [11] в 1944 году, $T_c \approx 2.269$, модель Бакстера-Ву решена Р.Бакстером и Ф.Ву [12] в 1973 году, $T_c \approx 2.269$, четырехкомпонентная модель Поттса решена Р.Поттсом [13] в 1952 году, $T_c \approx 0.91$.

Модели отличаются Гамильтонианами, топологией решеток и симметрией основного состояния. По этим признакам модели можно отнести к двум классам универсальности: Изинга, в который входит одноименная модель, и четырехкомпонентного Поттса, в который входят модели Бакстера-Ву и Поттса с q=4. Класс универсальности определяется набором критических показателей теплоемкости α, намагниченности β, восприимчивости γ, корреляционной длины ν и другими. Из гипотезы масштабной инвариантности [14, 15] следует ряд соотношений между критическими показателями [16—18]

$$2\beta + \gamma = 2 - \alpha = d\nu, \tag{2}$$

где d размерность пространства. Для класса универсальности Изинга $\alpha = 0, \beta = 1/8, \gamma = 7/4, \nu = 1$, для четырехкомпонентного Поттса $\alpha = 2/3, \beta = 1/12, \gamma = 7/6, \nu = 2/3$. Показатель ν отражает зависимость корреляционной длины при изменении линейного размера L решетки. Показатель ν и критическую температуру T_c в дальнейшем будем определять с помощью НС.

Классическим численным методом исследования фазовых переходов в решеточных спиновых моделях является метод Монте-Карло (МК). Генерация спиновых конфигураций методом МК происходит в соответствии с рекомендациями [19]. При генерации снимков конфигураций спинов подбираются параметры МК: i) время релаксации при "горячем" старте для минимизации систематической ошибки, ii) корреляционное время τ_{corr} при термодинамическом равновесии для минимизации статистической ошибки. Рассматриваются два класса алгоритмов: односпиновый алгоритм Метрополиса [20, 21] и кластерные алгоритмы Свендсена-Ванга [22] и Вольфа [23]. В алгоритме Метрополиса на каждом шаге принимается решение о выборе нового значения спина, а для кластерных – новой ориентации кластера спинов. Алгоритмы Свендсена-Ванга и Вольфа отличаются логикой формирования кластеров и эффективностью. Известно, что в критической области все алгоритмы страдают от критического замедления, при котором резко растет корреляционное время $\tau_{corr} \sim \min(L, \xi)^{d+z}$, где ξ корреляционная длина, z динамический критический показатель алгоритма. Наиболее эффективным является алгоритм Вольфа, но его реализация в модели Бакстера-Ву приводит к сдвигу T_c в низкотемпературную область, искажая результаты [24]. Был выбран алгоритм Метрополиса, наименее эффективный, но сохраняющий единый подход к генерации данных.

Вторая глава посвящена методам машинного обучения. Общая постановка задачи бинарной классификации в терминах МО – предсказать для каждого объекта $x_i \in X$ дискретное значение y_i, называемое меткой класса: $F(x_i) \to y_i$, где $y_i \in \{0, 1\}$, F решающая функция. Качество работы обученного классификатора измеряется метриками качества точности, полноты, f-меры и площади под ROC-кривой (AUCROC).

Решающей функцией F в диссертации является нейронная сеть. НС состоит из последовательных блоков, называемых слоями. Множество и порядок слоев в НС определяет её архитектуру. Каждый слой представляет собой дифференцируемую функцию. Слой имеет вход, выход и может состоять из линейных и нелинейных операторов. Основные составные элементы НС, используемые в диссертации, это полносвязные слои, сверточные слои, слои пулинга и функции активации. Предложены три архитектуры НС: полносвязная (FCNN), сверточная (CNN) и глубокая сверточная сеть ResNet [25].

Описаны основные составные элементы НС, механизмы настройки параметров сети, этапы обучения, валидации и тестирования, а также эвристики, влияющие на качество обучения. За настройку параметров отвечают механизмы обратного распространения ошибки [26, 27] и методы оптимизации на основе градиентного спуска [28]. Процесс обучения НС включает этапы разбиения данных, кодирования, контроль качества обучения и замер метрик. Эвристические подходы, позволяющие повысить метрики качества и ускорить процесс обучения: случайная инициализация весов НС, пакетная обработка данных, регуляризация входного потока данных (аугментация).

В **третьей главе** описан разработанный метод для анализа фазовых переходов в решеточных спиновых моделях. Решается задача бинарной классификации методом обучения с учителем. Для исследования критического поведения решеточных спиновых моделей, методом Монте-Карло алгоритмом Метрополиса генерируются спиновые конфигурации для разных линейных размеров решеток, по 1500 изображений на каждую температурную точку. Выходной слой каждой из архитектур НС состоит из двух нейронов для ферромагнитной и парамагнитной фаз. Тренируется НС под каждый размер L с функцией потерь бинарная кросс-энтропия, по графикам кривых обучения и валидации оценивается качество процесса обучения на нескольких эпохах. Для определения критических характеристик используется отложенная выборка, которую тестируют на весах НС после первой эпохи обучения.

Для НС под каждый линейный размер L строится функция вариации выхода $V(T)$ ферромагнитного нейрона f_i^T, функция определена в каждой температурной точке T

$$V(T) = \frac{1}{N}\sum_{i=1}^{N}\left(f_i^T\right)^2 - \left(\frac{1}{N}\sum_{i=1}^{N} f_i^T\right)^2, \tag{3}$$

где $i \in \{1, 2, ..N\}$ снимки конфигураций спинов при T. Исследуется зависимость $V(T)$ от линейного размера L. Для каждого значения L производится процедура аппроксимации $V(T)$ ненормированной Гауссовой кривой $V(T) \sim \exp\left(-(T - T_*)^2/2\sigma^2\right)$, извлекая параметры σ и T_*. Значения ширины σ кривой $V(T)$ аппроксимируются степенным законом $\sigma(L) \sim 1/L^{1/\nu}$ для извлечения критического показателя ν. Значения положения максимума T_* кривой $V(T)$ аппроксимируются законом Фердинанда-Фишера [29] $T_*(L) - T_c \sim 1/L^{1/\nu}$, при фиксированном значении ν для извлечения оценки критической температуры $\hat{T}_c$ при $L \to \infty$. Для моделей Изинга, Бакстера-Ву, четырехкомпонентного Поттса произведена численная оценка $\hat{T}_c$ и ν.

Остальная часть главы посвящена трем группам экспериментов. В первой, оценивается влияние размера обучающего набора данных на точность извлекаемых показателей, путем уменьшения обучающей выборки в два и в четыре раза. Во второй, исследуется влияние количества эпох обучения НС. При длительном обучении, НС лучше решает задачу классификации, что отражается на точности извлекаемых характеристик $\hat{T}_c$ и ν. В третьей, производится численная оценка критической температуры $\hat{T}_c$ и критического показателя ν через обучение с переносом знания с одной модели на другую.

В **четвертой главе** исследуется влияние способов кодирования входных данных на точность извлекаемых критических характеристик. В отличии от предыдущей главы, где НС обучается в пространстве спиновых конфигураций, применяются два других способа кодирования: в пространство корреляторов спинов и в пространство энергий связей.

Для кодирования в пространство корреляторов, формируется матрица размера $L \times L$, в ячейках которой находятся корреляторы спинов. При формирования матрицы используется априорное знание о топологии решетки. Каждый спин в узле взаимодействует с другими спинами на расстоянии $L/2$ по формуле

$$g_{x,y}(L/2) = \frac{1}{D}\sum_{d=1}^{D} s_{x,y}s_*, \tag{4}$$

где $g_{x,y}(L/2)$ значение коррелятора в узле (x,y), $s_{x,y}$ значение спина в том же узле, s_* значение спина стоящего на расстоянии $L/2$ от спина $s_{x,y}$ вдоль направления d, D – количество направлений, вдоль которых происходит взаимодействие.

Для кодирования в пространство энергий связей, формируются две матрицы размера $L \times L$ для хранения вертикальных $\{\sigma_V\}$ и горизонтальных $\{\sigma_H\}$ энергий связей. При формировании матриц используется априорное знание о взаимодействии между спинами:

$$\begin{aligned}
\{\sigma_H\} &= \{\sigma_{i,j}\sigma_{i,j+1}, \sigma_{i,j+1}\sigma_{i,j+2}, ..., \sigma_{i+L-1,j+L-1}\sigma_{i+L-1,j}\}, \\
\{\sigma_V\} &= \{\sigma_{i,j}\sigma_{i+1,j}, \sigma_{i+1,j}\sigma_{i+2,j}, ..., \sigma_{i+L-1,j+L-1}\sigma_{i,j+L-1}\}.
\end{aligned} \tag{5}$$

Производится полный цикл обучения НС по методу, описанному в главе 3 для извлечения оценок критической температуры T_c и критического показателя корреляционной длины ν с переносом обучения и без него. Обучение с переносом знания происходит по всем возможным комбинациям пар для моделей Изинга, Бакстера-Ву и четырехкомпонентного Поттса с кодированием в пространство энергий связей.

Положения выносимые на защиту

- Разработан метод анализа фазовых переходов в решеточных спиновых моделях на основе масштабирования функции вариации выхода нейронной сети решением задачи бинарной классификации с учителем. В методе описан процесс обучения НС, параметры, влияющие на точность извлечения критических характеристик, предложены рекомендованные значения гиперпараметров.

- Установлено, что функция вариации выхода НС, обученной по разработанному в диссертации методу, несет информацию о критическом показателе корреляционной длины ν и критической температуре $\hat{T}_c$ тестируемой решеточной спиновой модели. Относительная погрешность при извлечении равна 0.1-0.2% для $\hat{T}_c$, 1-3% для ν.

- Произведена численная оценка критического показателя корреляционной длины ν и критической температуры $\hat{T}_c$ для решеточных спиновых моделей Изинга $1/\nu = 1.02(1)$, $\hat{T}_c = 2.270(5)$; Бакстера-Ву $1/\nu = 1.49(2)$, $\hat{T}_c = 2.2691(1)$; четырехкомпонентного Поттса $1/\nu = 1.49(4)$, $\hat{T}_c = 0.9101(1)$.

- Метод численной оценки критических свойств решеточных спиновых моделей проверен с применением нескольких архитектур НС. Метод работает для полносвязной НС с одним скрытым слоем, неглубокой сверхточной НС и глубокой архитектуры ResNet с каскадом сверток. Точность работы каждой архитектуры зависит от свойств спиновой модели и гиперпараметров обучения.

- Метод численной оценки критических свойств чувствителен к гиперпараметрам обучения. Увеличение количества итераций обучения НС позволяет лучше решать задачу классификации, но ухудшает точность определения критического показателя ν. Увеличение размера обучающего набора данных не влияет на точность извлекаемых критических характеристик при достижении определенного уровня метрик качества классификации. Способ кодирования входных данных влияет на точность извлекаемых критических характеристик ν и $\hat{T}_c$.

- Разработан метод кодирования входных конфигураций спинов с помощью энергий связей. Произведена численная оценка критического показателя корреляционной длины ν и критической температуры $\hat{T}_c$ для решеточных спиновых моделей Изинга, Бакстера-Ву и Поттса с q=4 через обучение с переносом знания внутри и вне собственного класса универсальности. Результаты совпадают с точным решением, однако метод несистемный и чувствителен к параметрам решеточных спиновых моделей, архитектурам НС и гиперпараметрам обучения.

Личный вклад автора в разработку проблемы

Гипотезы, идеи и метод были разработаны совместно с научным руководителем. Этапы генерации данных методом Монте-Карло, обучение и тестирование НС, анализ результатов нейросетевого моделирования, валидация и сравнение с классическими методами были произведены лично соискателем. В публикациях вклад соискателя определяющий. Секционные доклады на конференциях были сделаны лично соискателем.

Общие выводы исследования

- Разработан метод для исследования фазовых переходов в решеточных спиновых моделях с использованием масштабирования функции вариации выхода НС $V(T)$. Обучение НС производилось решением задачи бинарной классификации обучения с учителем. Описан общий процесс обучения НС и параметры влияющие на точность извлекаемых критических характеристик. Применив метод к моделям Изинга, Бакстера-Ву, Поттса с q=4, были получены количественные оценки для критического показателя $1/\nu$ и критической температуры $\hat{T}_c$. Относительная погрешность работы метода для извлечения критической температуры $\hat{T}_c$ в пределах 0.2%, критического показателя ν в пределах 3%, что превосходит остальные подходы МО в решеточных спиновых моделях. Значения показателей ν и $\hat{T}_c$ в пределах статистической погрешности совпадают с точным аналитическим решением для этих моделей в классах универсальности Изинга и четырехкомпонентного Поттса.

- Анализ **архитектур** НС, от неглубоких к глубоким показал, что метод работает систематически. Однако, среди архитектур нельзя выбрать наилучшую, потому что их отличия в точности работы зависят от свойств исследуемых решеточных спиновых моделей, методов кодирования входных данных и способа исследования (с переносом обучения или без). При обучении на конфигурациях спинов более точные решения для $\hat{T}_c, \nu$ дают менее глубокие архитектуры FCNN и CNN, что согласуется с наблюдениями других исследователей [30]. При кодировании спиновых конфигураций в пространство энергий связей, архитектура FCNN сталкивается с недообучением, а точность работы CNN и ResNet не отличается друг от друга.

- Исследование влияния **количества эпох** при обучении показало, что при росте количества эпох, НС лучше решает задачу классификации, но хуже извлекает критический показатель ν, а на поздних эпохах превращается в идеальный классификатор без возможности извлекать ν. **Размер** набора данных при обучении НС может влиять на точность извлекаемых показателей до достижения определенного уровня метрик качества классификации. FCNN в модели Изинга при обучении на четверти выборки, сталкивается с недообучением по метрике AUCROC, что приводит ухудшению извлекаемого показателя ν.

- Произведена численная оценка критического показателя корреляционной длины ν и критической температуры $\hat{T}_c$ для решеточных спиновых моделей Изинга, Бакстера-Ву и Поттса с q=4 через обучение с переносом знания внутри и вне собственного класса универсальности. Перекрестное обучение работает менее систематически, чем тестирование на собственной модели. Критическая температура $\hat{T}_c$ извлекается систематически, для всех сочетаний обучения и тестирования РСМ, за исключением переноса знания с квадратной решетки на треугольную в пространстве спиновых конфигураций. Относительная погрешность составляет 0.1-1.5% в пространстве спиновых конфигураций, и 0.1-5% в пространстве энергий связей. Наибольшая относительная погрешность извлечения $\hat{T}_c$ при переносе знания с модели Поттса: 0.3-4% внутри класса универсальности, 7-16% между классами универсальности. Критический показатель корреляционной длины ν извлекается менее постоянно. В пространстве спиновых конфигураций перенос знания возможен только с класса универсальности Поттса на класс универсальности Изинга, и не возможен с квадратной решетки на треугольную. При использовании кодирования с помощью энергий связей, возможно перенести знание о фазовом переходе с модели Изинга на модель Бакстера-Ву с точностью как при собственном тестировании модели. При переходе от спиновых конфигураций в пространство энергий связей, НС не нужно “понимать” пространственное расположение спинов на решетке, проблема с которой сталкивались сети при переносе знания с квадратной решетки на треугольную. Кодирование спиновых конфигураций в пространство корреляторов и энергий связей происходит с внесением в НС дополнительной информации о топологии решетки и Гамильтониане взаимодействий в модели.

Апробация результатов

Публикации:

- Chertenkov V .I. Universality classes and machine learning / Shchur L.N. // Journal of Physics: Conference Series. 2021. N 1740. P. 1-5. (Scopus Q4)
- Chertenkov V .I. Deep machine learning investigation of phase transitions / Burovski E.A., Shchur L.N. // Lecture Notes in Computer Science. 2022. N 13708. P. 397-408. (Scopus Q2)
- Chertenkov V .I. Finite-size analysis in neural network classification of critical phenomena / Burovski E.A., Shchur L.N. // Physical Review E - Statistical, Nonlinear, and Soft Matter Physics. 2023. T. 108. N 3. P. 1-5. (Scopus Q1)
- Sukhoverhova D.D. Validity and limitations of supervised learning for phase transition research / Chertenkov V .I., Burovski E.A., Shchur L.N. // Lecture Notes in Computer Science. 2023. N 14389. P. 314-329. (Scopus Q2)

Секционные доклады:

- IV Международная конференция «Компьютерное моделирование в физике и не только», Россия, Москва, 12-16 октября 2020 г., “Universality classes and machine learning”.
- Международная конференция «Суперкомпьютерные дни в России», Россия, Москва, 26-27 сентября 2022 г., “Deep machine learning investigation of phase transitions”.
- Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В.Арменского, Россия, Москва, 27 февраля – 7 марта 2023 г., “Исследование спиновых моделей с помощью методов машинного обучения”.
- Национальный Суперкомпьютерный Форум НСКФ-2023, Россия, Переславль-Залесский, 28 ноября – 1 декабря 2023 г., “Influence of learning protocols on deep learning studies of phase transitions”.
- НИС Вычислительные среды, Москва, МИЭМ НИУ ВШЭ, 14 ноября 2023, “Unsupervised learning of phase transitions via modified anomaly detection with autoencoders”.
- Международная конференция «Суперкомпьютерные дни в России», Россия, Москва, 23-24 сентября 2024 г., “Supervised and Transfer Learning for Phase Transition Research”.

Зарегистрированное программное обеспечение:

- “Система исследования фазовых переходов в решеточных спиновых моделях”.

Список литературы

[1] Juan Carrasquilla и Roger G Melko. “Machine learning phases of matter”. В: *Nature Physics* 13.5 (2017), с. 431—434.

[2] Gordon Frank Newell. “Crystal statistics of a two-dimensional triangular Ising lattice”. В: *Physical Review* 79.5 (1950), с. 876.

[3] Kimihiko Fukushima и Kazumitsu Sakai. “Can a CNN trained on the Ising model detect the phase transition of the q-state Potts model?” В: *Progress of Theoretical and Experimental Physics* 2021.6 (2021), 061A01.

[4] Nicholas Walker и Ka-Ming Tam. “InfoCGAN classification of 2D square Ising configurations”. В: *Machine Learning: Science and Technology* 2.2 (2020), с. 025001.

[5] Constantia Alexandrou и др. “The critical temperature of the 2D-Ising model through deep learning autoencoders”. В: *The European Physical Journal B* 93 (2020), с. 1—15.

[6] Wanzhou Zhang, Jiayu Liu и Tzu-Chieh Wei. “Machine learning of phase transitions in the percolation and X Y models”. В: *Physical Review E* 99.3 (2019), с. 032142.

[7] Yusuke Miyajima и др. “Machine learning detection of Berezinskii-Kosterlitz-Thouless transitions in q-state clock models”. В: *Physical Review B* 104.7 (2021), с. 075114.

[8] Kenta Shiina и др. “Machine-learning studies on spin models”. В: *Scientific reports* 10.1 (2020), с. 2177.

[9] Dimitrios Bachtis, Gert Aarts и Biagio Lucini. “Mapping distinct phase transitions to a neural network”. В: *Physical Review E* 102.5 (2020), с. 053306.

[10] Ernst Ising. “Beitrag zur theorie des ferro-und paramagnetismus”. Дис. ... док. Grefe & Tiedemann Hamburg, Germany, 1924.

[11] Lars Onsager. “Crystal statistics. I. A two-dimensional model with an order-disorder transition”. В: *Physical Review* 65.3-4 (1944), с. 117.

[12] RJ Baxter и FY Wu. “Exact solution of an Ising model with three-spin interactions on a triangular lattice”. В: *Physical Review Letters* 31.21 (1973), с. 1294.

[13] Renfrey Burnard Potts. “Some generalized order-disorder transformations”. В: *Mathematical proceedings of the cambridge philosophical society.* Т. 48. 1. Cambridge University Press. 1952, с. 106—109.

[14] Leo P Kadanoff. “Scaling laws for Ising models near T c”. В: *Physics Physique Fizika* 2.6 (1966), с. 263.

[15] АЗ Паташинский и ВЛ Покровский. “О поведении упорядочивающихся систем вблизи точек фазового перехода”. В: *ЖЭТФ* 50.2 (1966), с. 439—447.

[16] Benjamin Widom. “Equation of state in the neighborhood of the critical point”. В: *The Journal of Chemical Physics* 43.11 (1965), с. 3898—3905.

[17] Michael E Fisher. “The theory of equilibrium critical phenomena”. В: *Reports on progress in physics* 30.2 (1967), с. 615.

[18] Leo P Kadanoff и др. “Static phenomena near critical points: theory and experiment”. В: *Reviews of Modern Physics* 39.2 (1967), с. 395.

[19] Alan Sokal. “Monte Carlo methods in statistical mechanics: foundations and new algorithms”. В: *Functional integration: Basics and applications*. Springer, 1997, с. 131—192.

[20] W Keith Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. В: (1970).

[21] Nicholas Metropolis и др. “Equation of state calculations by fast computing machines”. В: *The journal of chemical physics* 21.6 (1953), с. 1087—1092.

[22] Robert H Swendsen и Jian-Sheng Wang. “Nonuniversal critical dynamics in Monte Carlo simulations”. В: *Physical review letters* 58.2 (1987), с. 86.

[23] Ulli Wolff. “Collective Monte Carlo updating for spin systems”. В: *Physical Review Letters* 62.4 (1989), с. 361.

[24] Lev N Shchur и Wolfhard Janke. “Critical amplitude ratios of the Baxter–Wu model”. В: *Nuclear Physics B* 840.3 (2010), с. 491—512.

[25] Kaiming He и др. “Deep residual learning for image recognition”. В: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, с. 770—778.

[26] Paul Werbos. “Beyond regression: New tools for prediction and analysis in the behavioral sciences”. В: *PhD thesis, Committee on Applied Mathematics, Harvard University, Cambridge, MA* (1974).

[27] David E Rumelhart, Geoffrey E Hinton и Ronald J Williams. “Learning representations by back-propagating errors”. В: *nature* 323.6088 (1986), с. 533—536.

[28] Diederik P Kingma и Jimmy Ba. “Adam: A method for stochastic optimization”. В: *arXiv preprint arXiv:1412.6980* (2014).

[29] Arthur E Ferdinand и Michael E Fisher. “Bounded and inhomogeneous Ising models. I. Specific-heat anomaly of a finite lattice”. В: *Physical Review* 185.2 (1969), с. 832.

[30] Alan Morningstar и Roger G Melko. “Deep learning the ising model near criticality”. В: *Journal of Machine Learning Research* 18.163 (2018), с. 1—17.