

УДК 004.032.26
MSC2020 68T10

© А. О. Король^{1,2}; В. Ю. Капитан^{1,2}

Нейронная сеть для определения температуры Кюри двумерной модели Изинга

В работе описывается метод определения критической точки фазового перехода II рода двумерной модели Изинга на квадратной решётке с помощью свёрточной нейронной сети. Данные для обучения нейронной сети рассчитаны алгоритмом Метрополиса. Нейронная сеть обучалась на данных, соответствующих низкотемпературной фазе — ферромагнитной и высокотемпературной — парамагнитной соответственно. После обучения, свёрточная нейронная сеть анализировала входные данные из всего температурного диапазона от 0.1 до 5.0 (в безразмерных величинах) и определяла температуру Кюри T_c . Точность полученных результатов оценивалась относительно решения Онсагера для плоской решетки спинов Изинга.

Ключевые слова: *модель Изинга, температура Кюри, метод Монте-Карло, свёрточная нейронная сеть.*

DOI: <https://doi.org/10.47910/FEMJ202105>

Введение

Историю развития нейронных сетей и их приложений в физических науках можно проследить с середины XX века [1], начиная с внедрения методов на основе байесовских и простых искусственных нейронных сетей, но именно последнее десятилетие стало свидетелем смены парадигмы в способах решения фундаментальных и прикладных задач в физике, материаловедении и других смежных областях. Способность решать проблемы оптимизации жизненно важна в современном мире, поскольку постоянно возрастает сложность изучаемых систем и законов взаимодействий между их элементами. В связи с чем такие классы задач принципиально трудно решать, поскольку они требуют поиска потенциальных решений на огромных пространствах возможных состояний.

¹ Дальневосточный федеральный университет, 690950, г. Владивосток, ул. Суханова, 8.

² Институт прикладной математики ДВО РАН, 690041, г. Владивосток, ул. Радио, 7.

Электронная почта: kapitan.vyu@dvfu.ru (В. Ю. Капитан).

В статистической физике численное моделирование успешно используется при изучении фазовых переходов и критических явлений. Алгоритм Метрополиса как метод Монте-Карло — это общий метод, использующийся для исследования термодинамических свойств веществ, состоящих из взаимодействующих частиц. Алгоритм состоит из двух основных частей: равновероятного семплирования пространства состояний и оценки физических величин на основе выборок [2]. Для физики конденсированного состояния характерен рост сложности анализа исследуемых конфигураций в зависимости от числа частиц, однако уже разрабатываются методы машинного обучения, позволяющие классифицировать и интерпретировать сложные наборы данных и которые, уже представляют собой новый класс исследовательских вычислительных инструментов. Применение машинного обучения в статистической физике началось сравнительно недавно, но стремительно развивается. Нейронные сети стали популярным инструментом благодаря высокой скорости их обучения и точности прогнозов, а также простоте использования программных библиотек. На данный момент существует два основных подхода к применению нейронных сетей для исследования спиновых систем.

Первый является альтернативой исследованию физических величин с помощью методов Монте-Карло, когда изучение спиновых конфигураций состоит в классификации неупорядоченных и упорядоченных состояний с использованием алгоритмов машинного обучения, таких как свёрточная нейронная сеть. Использование такого подхода сводит проблему определения фаз и фазового перехода к проблеме классификации изображений, по сути, к основной проблемной области, в которой используются нейронные сети [3, 4]. Второй подход заключается в использовании нейронных сетей для предсказания спиновых конфигураций с наименьшей энергией, то есть основных состояний, применяя, например, авторегрессионную нейронную сеть или машину Больцмана.

Основным преимуществом нейронных сетей является способность обобщать различные задачи. Например, обучив сеть на конфигурациях для ферромагнитной модели Изинга, возможно получить критические параметры для решетки с другой структурой по имеющимся для неё данным (т.н. трансферное обучение), рассчитанным в результате моделирования методом Монте-Карло. Кроме того, применение нейронных сетей дает выигрыш во времени и точности по сравнению с МК-моделированием. Таким образом, машинное обучение начинает дополнять традиционно применяемые в научных исследованиях методы статистического анализа.

1. Описание модели и методов

В данном исследовании был применен алгоритм Метрополиса для Монте-Карло моделирования при вычислении температуры Кюри в рамках двумерной модели Изинга, а также для генерации входных данных для нейронной сети и затем было проведено сравнение с результатами, полученными после обучения свёрточной нейронной сети.

1.1. Модель Изинга

Модель Изинга — это простейшая из математических моделей статистической физики, применяемая для исследования фазовых переходов и критических точек, имеющая точное решение: $\frac{T_c}{J} = \frac{2}{\ln(1+\sqrt{2})} = 2.269$, с которым сравнивались получаемые другими методами данные. Поэтому данная модель является прекрасным базисом для начала работы с новыми методами изучения магнитных систем — нейронными сетями.

В работе использовался простейший гамильтониан для квадратной решетки спинов Изинга с четырьмя ближайшими соседями и замкнутыми граничными условиями. Математическая модель представляет собой совокупность дискретных переменных (значения магнитных моментов атомных спинов), которые могут принимать одно из двух значений $S_i = \pm 1$, соответствующих одному из двух состояний. Система спинов Изинга имеет размер $N = L \times L$. Гамильтониан имеет вид

$$H = -J \sum_{\langle ij \rangle} S_i S_j, \quad (1)$$

где $J = 1$ — константа ферромагнитного взаимодействия.

1.2. Алгоритм Метрополиса

Монте-Карло моделирование с помощью алгоритма Метрополиса используется во многих областях науки, поскольку имеет ряд преимуществ:

- 1) алгоритм позволяет корректно учитывать температуру;
- 2) позволяет исследовать наночастицы и их ансамбли вплоть до атомов, что дает возможность подробного изучения их структуры.

Из недостатков — в точке, где температура близка к температуре фазового перехода, алгоритм замедляется и число шагов, необходимое для приведения системы в состояние термодинамического равновесия, возрастает по экспоненциальному закону. Для разрешения этой проблемы используются другие методы Монте-Карло [5], но в пользу большей точности жертвуем скоростью вычислений, поэтому сейчас многие ученые начинают рассматривать альтернативу — нейронные сети.

Алгоритм Метрополиса используется для нахождения глобального минимума. Основная идея заключается в равномерном семплировании пространства состояний с заданной вероятностью распределения. На каждой итерации выборки конфигурация системы меняется за счет изменения ориентации случайно выбранного спина. Данный алгоритм был уже подробно рассмотрен авторами в прошлых работах [5, 6], поэтому его детального описания в настоящей работе о применении нейронных сетей не приводится.

В нашем исследовании были использованы конфигурации спиновых систем, полученные при разных температурах, для обучения и последующего анализа в нейронной сети. А также результаты МК-моделирования сравниваются с результатами работы нейронной сети и точным решением Онсагера для модели Изинга с четырьмя ближайшими соседями.

1.3. Нейронные сети

Применение нейронных сетей в компьютерном моделировании приобретает все большую популярность по причине точности, скорости работы и наличия множества программных библиотек.

Нейронные сети для глубокого обучения представляют собой взвешенный направленный граф, узлы которого (нейроны в нашем случае) расположены слой за слоем. Каждый нейрон предыдущего слоя связан с нейронами последующего. Эти связи между ними имеют определенные веса, которые являются основной частью нейронных сетей: от весов связей зависит предсказание нейронной сети. Каждый нейрон имеет активационную функцию, которая отвечает за вес связи на его выходе. Таким образом, каждый слой архитектуры анализирует значения нейронов входного слоя, а активационные функции преобразуют значения весов между нейронами.

Нейронные сети обычно обучаются с учителем, что предусматривает наличие обучающих данных, которые состоят из примеров и соответствующих им значений: меток, классов, параметров. Обучение сети может происходить двумя методами: прямого и обратного распространения. Метод обратного распространения более предпочтителен, поскольку корректирует веса связей в соответствии с величиной ошибки [7].

1.3.1. Структура нейронной сети и процесс обучения

В настоящей работе было продемонстрировано, что современные архитектуры машинного обучения могут предоставить новые подходы к изучению физических систем в рамках моделей статистической физики. Для этого была использована библиотека TensorFlow для построения свёрточной нейронной сети [8].

Наиболее точные результаты анализа сегодня демонстрируют нейронные сети на основе сверточной архитектуры. Когда нейроны связаны только с несколькими нейронами следующего слоя, слой называется сверточным. Сверточный слой действует как фильтр, который отсеивает наименее информативные части входных данных.

Созданная в ходе нашего исследования сверточная нейронная сеть состоит из нескольких слоев (см. рис. (1) и описание каждого слоя ниже).

1) Входной слой

Конфигурация модели Изинга (входные данные), каждому из нейронов (спинов) которой присвоен начальный случайный вес.

2) Сверточный слой с фильтром 3×3

У каждого слоя есть фильтры — матрицы со значениями весов. При движении фильтра по матрице предыдущего слоя происходит умножение каждого элемента фильтра на значение нейрона, далее значения суммируются и записываются в карту признаков:

$$x_j^l = f\left(\sum_i x_i^{l-1} * k_j^l + b_j^l\right), \quad (2)$$

При тестировании нейронной сети нейроны уже не выбрасываются, но выход каждого нейрона умножается на $(1 - p)$.

5) Выравнивающий слой

Слой, который формирует вектор нейронов из матрицы нейронов.

6) Полносвязный слой

Полносвязные слои используются для классификации. Все слои, идущие до полносвязных, используются для выделения различных признаков, которые подаются на вход классификатору.

$$x_j^k = f\left(\sum_{i=1}^m w_{ij}x_i^{k-1}\right), \quad (4)$$

где x_j^k — выходной сигнал для нейрона j на слое k , x_i^{k-1} — входной сигнал, w_{ij} — весовой коэффициент, $f()$ — функция активации.

На каждом слое применяется функция активации *ReLU*: $A(x) = \max(0, x)$, которая возвращает значение x , если x положительно, в противном случае возвращает ноль. Таким образом, *ReLU* работает как аппроксиматор.

На последнем слое применяется функция *Softmax*. Пусть M — количество нейронов в уровне с функцией активации *Softmax*, s_j — сигнал активации на j -ом нейроне. Тогда выходом на j -ом нейроне с функцией активации *Softmax* будет $f_j(s_1, \dots, s_M)$:

$$f_j(s_1, \dots, s_M) = \frac{e^{s_j}}{\sum_{k=1}^M e^{s_k}}, j = 1, \dots, M. \quad (5)$$

2. Результаты и обсуждение

В ходе исследования были проверены различные наборы входных данных нейронной сети, полученные при различных параметрах работы алгоритма Метрополиса при изучении систем 10×10 и 20×20 спинов для выбора оптимальных параметров моделирования, которые будут применяться при изучении более сложных спиновых систем. Проведен сравнительный анализ с результатами МК-моделирования.

На рис. 2 (а,б), представлены результаты применения нейронных сетей к вычислению критической точки T_c . На первом этапе сеть обучалась на спиновых конфигурациях, полученных в ходе МК-моделирования со следующими параметрами: размер системы — 10×10 , $T = 0.1 \dots 5.0$ с шагом 0.01 — число МК-шагов для термализации — 10000, число МК-шагов для расчета термодинамических средних в алгоритме Метрополиса — 10000, размер выборки конфигураций для обучения сети — 50 на один шаг по температуре, результаты представлены на рис. 2 (а). При этом в нейронной сети в качестве примера функции активации была визуализована сигмоида $f(s) = \frac{1}{1 + \exp(-s)}$, на которой можно увидеть, как обучалась сеть на данных, полученных ниже и выше области критической температуры соответственно.

На рис.2 (б) показан результат анализа входных данных уже без предварительного обучения сети, при этом подавались более “грубые” данные: шаг по $T = 0.1$, а число МК-шагов было равно 1000 при прочих одинаковых параметрах. Как видно на графиках, оценка T_c для алгоритма Метрополиса ухудшилась с уменьшением числа МК-шагов и увеличением шага по температуре. На результаты работы нейронной сети это практически не повлияло.

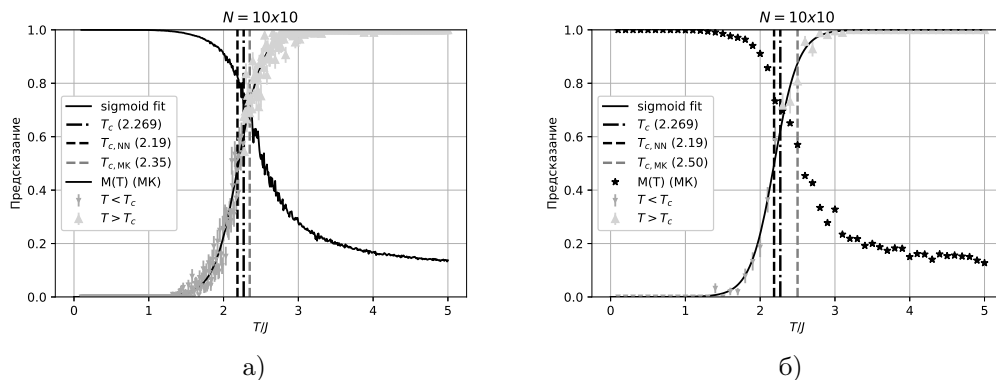


Рис. 2. Результаты расчетов T_c различными методами

По аналогии с описанием, представленным в п.1.3.1, в ходе МК-моделирования были использованы различные шаги по T и изменялось число МК-шагов. Увеличение размера системы положительно повлияло на результаты МК моделирования при расчетах T_c . За счет уменьшения влияния размерного эффекта $T_c = 2.29$. Увеличение размера системы существенно не отразилось на результатах работы нейронной сети, точность предсказанного значения критической температуры (по сравнению с описанных выше случаем) не изменилась и составила 2.6%. Существенной разницы между обучением на точных данных и проверкой на более грубых (и наоборот) не зафиксировано. В обоих случаях получены примерно одинаковые результаты. Увеличение размера выборки как для обучения, так и для тестирования не повлияло на точность получаемых значений температуры Кюри.

С точки зрения дальнейшего практического использования нейронных сетей был зафиксирован заслуживающий внимания результат. Нейронная сеть обучалась на “грубых” данных, аналогичных для случая, зафиксированного на рис.2 (б), с числом выборки 50. При этом тестирование проводилось на выборке, увеличенной до 100 образцов; другие параметры МК-моделирования не менялись, см. рис.3 (а, б).

Как показано на рис. 3 (б), при анализе большей по размеру выборки состояний по сравнению с размером выборки, примененной для обучения (рис. 3 (а)), получены более точные данные и сеть дополнительно “дообучилась” на таких данных.

В среднем можно отметить большую точность получаемых результатов по сравнению с Монте-Карло моделированием. Планируется дальнейшее изучение возможности применения свёрточных нейронных сетей в компьютерном моделировании: не только анализ более “грубых” входных данных, но и анализ других решеток, а так-

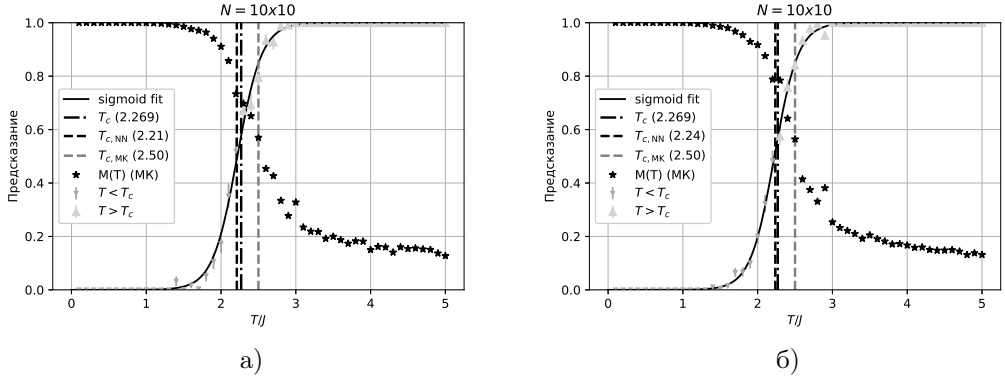


Рис. 3. Результаты расчетов T_c различными методами

же и проверка возможности обучения сети на простой модели для последующего анализа более сложных моделей статистической физики.

Заключение

В работе было рассмотрено применение нейронных сетей для определения критической температуры фазового перехода II рода в сравнении с известными решениями. Успешно была показана возможность применения свёрточных нейронных сетей к таким задачам, сведенным к задаче классификации спиновых состояний при разных температурах. Показана зависимость точности обучения сети и ее последующего применения от числа Монте-Карло шагов и размера выборки, приведено сравнение с алгоритмом Метрополиса. Рассмотрены системы различных размеров и влияние размерного эффекта на точность получаемых результатов. Также авторы отмечают такую особенность результатов, полученных с помощью нейронных сетей для определения T_c : если расчет производится с помощью алгоритма Метрополиса, то всегда $T_c^{MM} \geq T_c^{Exact}$. В свою очередь, при расчётах, проведенных с помощью свёрточных нейронных сетей $T_c^{NN} \leq T_c^{Exact}$. Причины такого поведения — предмет будущих исследований, в ходе которых планируется применить нейронные сети к изучению более сложных моделей и решеток.

Список литературы

- [1] A.M. Turing, “Computing Machinery and Intelligence”, *Mind*, **59**:236 (1950), 433–460.
- [2] А. Г. Макаров и др., “К численному расчету фрустраций в модели Изинга”, *Письма в Журнал экспериментальной и теоретической физики*, **110**:10 (2019), 700–705.
- [3] J. Carrasquilla and R.G. Melko, “Machine learning phases of matter”, *Nature Physics*, **13**:5 (2017), 431–434.
- [4] S. Kenta, et al., “Machine-Learning Studies on Spin Models”, *Scientific Reports*, **10**:1 (2020).

- [5] К. В. Шаповалова и др., “Методы канонического и мультиканонического семплирования пространства состояний векторных моделей”, *Дальневосточный математический журнал*, **17**:1 (2017), 124–130.
- [6] В. Ю. Капитан и др., “Термодинамические свойства систем спинов Гейзенберга на квадратной решетке с взаимодействием Дзялошинского–Мория”, *Дальневосточный математический журнал*, **20**:1 (2020), 63–73.
- [7] I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, MIT press, 2016.
- [8] M. Abadi, et al., “TensorFlow: Large-scale machine learning on heterogeneous systems”, 2015, TensorFlow.

Поступила в редакцию
11 мая 2020 г.

Работа выполнена в рамках Государственного задания Министерства науки и высшего образования РФ № 0657-2020-0005 от 17.03.2020.

*Korol A. O.*¹, *Kapitan V. Yu.*^{1,2} Neural Network for Prediction of Curie Temperature of Two-Dimensional Ising Model. *Far Eastern Mathematical Journal*. 2021. V. 21. No 1. P. 51–60.

¹ Far Eastern Federal University, Russia

² Institute for Applied Mathematics, Far Eastern Branch, Russian Academy of Sciences

ABSTRACT

The authors describe a method for determining the critical point of a second-order phase transitions using a convolutional neural network based on the Ising model on a square lattice. Data for training were obtained using Metropolis algorithm for different temperatures. The neural network was trained on the data corresponding to the low-temperature phase, that is a ferromagnetic one and high-temperature phase, that is a paramagnetic one, respectively. After training, the neural network analyzed input data from the entire temperature range: from 0.1 to 5.0 (in dimensionless units) and determined the Curie temperature T_c . The accuracy of the obtained results was estimated relative to the Onsager solution for a flat lattice of Ising spins.

Key words: *Ising model, Curie temperature, Monte Carlo method, Convolutional neural network.*

References

- [1] A.M. Turing, “Computing Machinery and Intelligence”, *Mind*, **59**:236 (1950), 433–460.
- [2] A. G. Makarov i dr., “K chislennomu raschetu frustratsii v modeli Izinga”, *Pis'ma v Zhurnal eksperimental'noi i teoreticheskoi fiziki*, **110**:10 (2019), 700–705.
- [3] J. Carrasquilla and R.G. Melko, “Machine learning phases of matter”, *Nature Physics*, **13**:5 (2017), 431–434.
- [4] S. Kenta, et al., “Machine-Learning Studies on Spin Models”, *Scientific Reports*, **10**:1 (2020).

- [5] K. V. Shapovalova i dr., “Metody kanonicheskogo i mul'tikanonicheskogo simplirovaniia prostranstva sostoianii vektornykh modelei”, *Dal'nevostochnyi matematicheskii zhurnal*, **17**:1 (2017), 124–130.
- [6] V. Iu. Kapitan i dr., “Termodinamicheskie svoistva sistem spinov Geizenberga na kvadratnoi reshetke s vzaimodeistviem Dzialoshinskogo–Moriia”, *Dal'nevostochnyi matematicheskii zhurnal*, **20**:1 (2020), 63–73.
- [7] I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, MIT press, 2016.
- [8] M. Abadi, et al., “TensorFlow: Large-scale machine learning on heterogeneous systems”, 2015, TensorFlow.