

УДК 004.032.26, 536.911

НЕЙРОСЕТЕВОЙ МЕТОД РАСЧЕТА ТОЧКИ КЮРИ ДВУМЕРНОЙ МОДЕЛИ ИЗИНГА

А.О. Король^{1,2}, В.Ю. Капитан^{1,2}

¹ Дальневосточный федеральный университет, 690922, г. Владивосток, остров Русский, п. Аякс,
10, Кампус ДВФУ, Россия

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

В работе описывается метод определения критической точки фазового перехода II рода с помощью свёрточной нейронной сети на примере модели Изинга на квадратной решётке. Данные для обучения и анализа получены с помощью Монте-Карло моделирования для разных температур. Нейронная сеть обучалась на данных, соответствующих низкотемпературной фазе – ферромагнитной и высокотемпературной – парамагнитной, соответственно. После обучения нейронная сеть анализировала входные данные из всего температурного диапазона: от 0.1 до 5.0 (в безразмерных величинах) и определяла точку Кюри T_c .

Ключевые слова:

модель Изинга, точка Кюри, метод Монте-Карло, свёрточная нейронная сеть.

Благодарности

Работа выполнена при финансовой поддержке государственного задания Министерства науки и высшего образования РФ №075-00400-19-01.

THE NEURAL NETWORK METHOD FOR CALCULATING THE CURIE POINT OF A TWO-DIMENSIONAL ISING MODEL

A.O. Korol^{1,2}, V.Yu. Kapitan^{1,2}

¹ Far Eastern Federal University, FEFU Campus, 10 Ajax Bay, Russky Island, Vladivostok, 690922,
Russia

² Institute of Applied Mathematics, Far Eastern Branch, Russian Academy of Science, 7 Radio St.,
Vladivostok, 690041, Russia

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 and analysis were obtained using Monte Carlo simulations 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 point T_c .

Keywords

Ising model, Curie point, Monte Carlo method, Convolutional neural network.

Acknowledgements

This work was financially supported by the state task of the Ministry of Science and Higher Education of Russia No. 075-00400-19-01.

Сведения об авторах / Information about the Authors

Алена Олеговна Король, бакалавр, Кафедра компьютерных систем, Школа естественных наук, Дальневосточный федеральный университет, 690922, г. Владивосток, остров Русский, п. Аякс, 10, Кампус ДВФУ, Россия

Лаборант-исследователь, Институт прикладной математики ДВО РАН, 690041, г. Владивосток, ул. Радио, 7, Россия

Alena O. Korol, bachelor, Department of Computer Systems, School of Natural Sciences, Far Eastern Federal University, FEFU Campus, 10 Ajax Bay, Russky Island, Vladivostok, 690922, Russia

Laboratory researcher, Institute of Applied Mathematics, Far Eastern Branch, Russian Academy of Science, 7 Radio St., Vladivostok, 690041, Russia

korol.ao@students.dvfu.ru

Виталий Юрьевич Капитан, кандидат физико-математических наук, доцент, Кафедра компьютерных систем, Школа естественных наук, Дальневосточный федеральный университет, 690922, г. Владивосток, остров Русский, п. Аякс, 10, Кампус ДВФУ, Россия

Старший научный сотрудник, Институт прикладной математики ДВО РАН, 690041, г. Владивосток, ул. Радио, 7, Россия

Vitalii Yu. Kapitan, Candidate of Science (Physics and Mathematics), Associate professor, Department of Computer Systems, School of Natural Sciences, Far Eastern Federal University, FEFU Campus, 10 Ajax Bay, Russky Island, Vladivostok, 690922, Russia

Senior Researcher, Institute of Applied Mathematics, Far Eastern Branch, Russian Academy of Science, 7 Radio St., Vladivostok, 690041, Russia

kapitan.vyu@dvfu.ru

Введение

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 [1,2]. Алгоритм состоит из двух основных частей: равновероятного семплирования пространства состояний и оценки физических величин на

основе выборок [3]. Благодаря развитию суперкомпьютерных вычислительных систем [4], алгоритмы продолжают активно развиваться и совершенствоваться [5]. При этом для физики конденсированного состояния характерна постоянно возрастающая сложность анализа исследуемых конфигураций в зависимости от числа частиц, однако уже разработаны методы машинного обучения, позволяющие классифицировать и анализировать сложные наборы данных и как следствие осуществлять разработку новых исследовательских методов [6].

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. 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 нее данным, рассчитанным в результате моделирования методом Монте-Карло [6,7]. Поэтому машинное обучение начинает активно 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 модели Изинга, а также для генерации входных данных для нейронной сети и затем было проведено сравнение с результатами, полученными после обучения свёрточной нейронной сети. Все величины в работе приведены в безразмерных величинах.

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.

В работе использовался 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 $N = L \times L$ и гамильтониан имеет вид:

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

где $J=1$ – константа ферромагнитного взаимодействия [8].

1.2. Алгоритм Метрополиса

Монте-Карло моделирование с помощью алгоритма Метрополиса используется во многих областях и позволяет исследовать термодинамические свойства веществ, состоящих из взаимодействующих частиц. Основная идея алгоритма заключается в равномерном семплировании пространства состояний с заданной вероятностью распределения. На каждой итерации выборки конфигурация системы меняется за счет изменения ориентации случайно выбранного спина; алгоритм был детально рассмотрен авторами ранее [3,8]. Однако данный метод не лишен недостатков, так в точке, где температура близка к температуре фазового перехода, он замедляется и число шагов, необходимое для приведения системы в состояние термодинамического равновесия, возрастает по экспоненциальному закону. Соответственно для разрешения этой проблемы используются другие методы Монте-Карло, но в пользу большей точности, мы жертвуем скоростью вычислений, поэтому сейчас многие ученые начинают рассматривать альтернативу – нейронные сети.

В данном исследовании были использованы конфигурации спиновых систем, полученные при разных температурах с помощью алгоритма Метрополиса, для обучения и последующего анализа в свёрточной нейронной сети. А также результаты МК-моделирования сравнивались 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 в силу своей точности, скорости 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 [9].

1.3.1. Структура нейронной сети и процесс обучения

В работе было продемонстрировано, что современные методы 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 [10].

В данный момент наиболее точные результаты анализа демонстрируют нейронные сети на основе свёрточной архитектуры. Когда нейроны связаны только с несколькими нейронами следующего слоя, слой называется свёрточным. Свёрточный слой действует как фильтр, который отсеивает наименее информативные части входных данных.

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

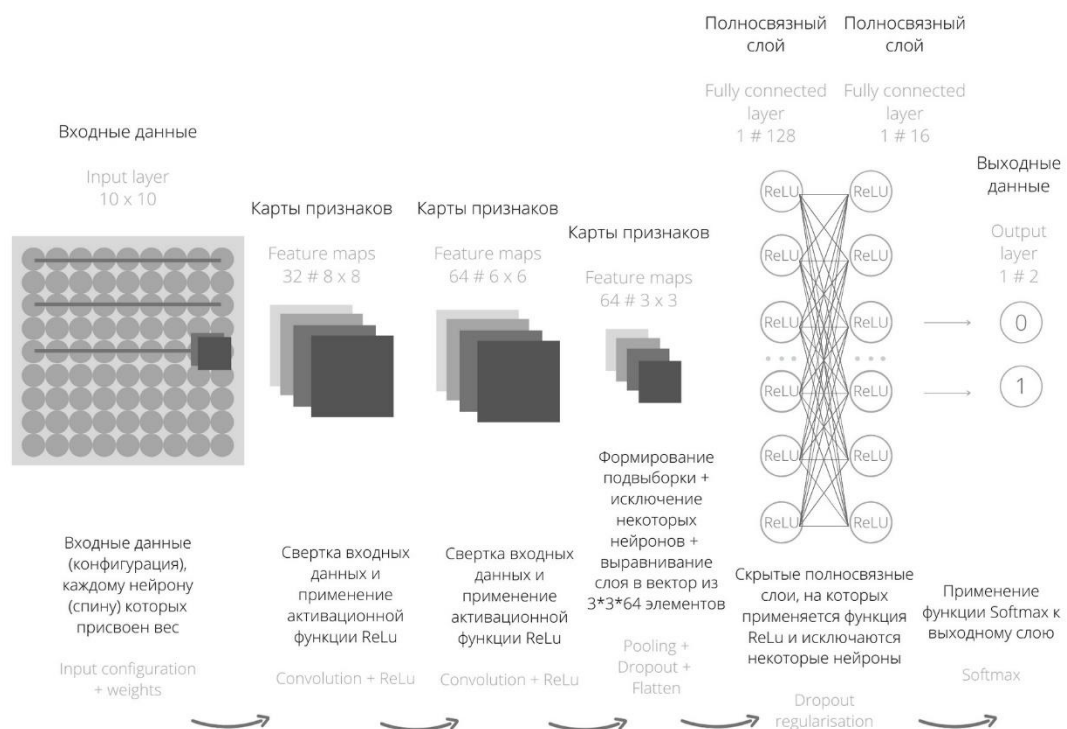


Рис. 1. Структура свёрточной нейронной сети

1) Входной слой

Входные данные (конфигурация модели Изинга), каждой из нейронов (спинов) которой присвоен начальный случайный вес.

2) Свёрточный 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:

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

где x_j^l – карта признаков j (выходной сигнал на уровне l), $f()$ – функция активации, применяемая поэлементно, b_j^l – пороговое значение (сдвиг) для карты признаков j , k_j^l – ядро свёртки с номером j , x_i^{l-1} – карты признаков предыдущего уровня, $*$ – операция свёртки.

3) Слой подвыборки

Слой подвыборки снижает размерность поступивших на него данных путем выбора максимальных значений из «окна подвыборки» (pool). В нашем случае $pool_{size} = 2$:

$$x^l = f(a^l subsample(x^{l-1}) + b^l), \quad (3)$$

где x^l – выход уровня l , $f()$ – функция активации, применяемая поэлементно, a^l , b^l – коэффициенты, $subsample$ – операция выборки локальных максимальных значений, т. е. максимальных значений в каждом окне *max-pooling*.

4) Исключающий слой

Для того, чтобы исключить возможность переобучения сети, используется слой *Dropout*, который исключает нейроны из сети с заданной вероятностью p . Исключенные нейроны не вносят вклад в дальнейшее обучение сети. Таким образом, получаем одну из возможных 2^N структур связей между нейронами, где N – суммарное число нейронов.

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.

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

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

На рис. 2 (а,б), представлены результаты применения свёрточной нейронной сети к вычислению критической точки T_c в сравнении с точным решением Онсагера и результатами МК-моделирования. 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 предварительного уравнивания системы: 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 качестве примера функции активации в нейронной сети, была визуализована сигмоида: $f(s) = 1/(1 + \exp(-s))$, на которой можно увидеть, как обучалась сеть на данных, полученных ниже и выше области критической температуры, соответственно. На рис. 2 (б) показан результат анализа более «грубых» входных данных: шаг по $T = 0.1$, а число МК-шагов было сокращено до 1000, при прочих одинаковых параметрах моделирования. Как видно из графиков, оценка T_c для алгоритма Метрополиса ухудшилась с уменьшением числа МК-шагов и увеличением шага по температуре. На результаты работы нейронной сети это не повлияло.

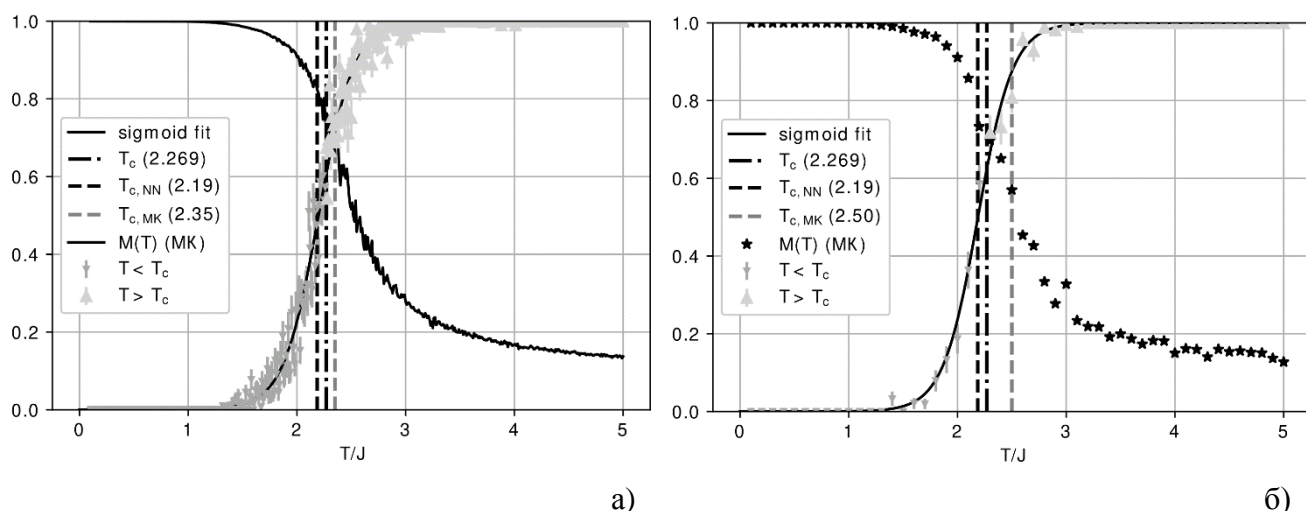


Рис. 2. Результаты расчетов T_c различными методами для системы 10×10 спинов Изинга:

а) шаг по $T = 0.01$, число МК-шагов равно 10000, б) шаг по $T = 0.1$, число МК-шагов равно 1000

Также было проверено влияние размера системы на точность получаемых результатов, на примере системы 20×20 . В ходе моделирования также были использованы различные параметры для МК-моделирования, аналогично описанию, приведенному выше: изменялась величина шага по T и число МК-шагов. Полученные результаты в целом аналогичны приведенным выше. Стоит отметить, что увеличение размера системы положительно повлияло на результаты МК-моделирования при расчетах T_c : $T_c = 2.29$, за счет уменьшения влияния размерного эффекта, при этом на результатах работы нейронной сети, увеличение размера системы, существенно не отразилось, точность предсказанного значения критической температуры, по сравнению со случаем описанным выше, в среднем не изменилась и составила $T_c = 2.21$, а в некоторых численных экспериментах даже ухудшилась, т.к. обучение сети строится на вероятностном подходе. Также можно отметить, что авторы не зафиксировали существенной разницы между обучением сначала на точных данных, а потом проверки более грубых, так и наоборот. Результаты в обоих случаях были примерно одинаковы. Увеличение размера выборки и для обучения, и для тестирования, также сильно не повлияли на точность.

Однако, заслуживающий внимания результат, с точки зрения дальнейшего использования нейронных сетей при анализе спиновых систем, был зафиксирован, когда сеть обучалась на «грубых» данных, аналогичных для случая описанного на рис. 2 (б) с числом выборки 50. А вот тестирование проводилось на увеличенной до 100 образцов выборке. При этом другие параметры МК-моделирования не менялись, см. рис. 3 (а,б).

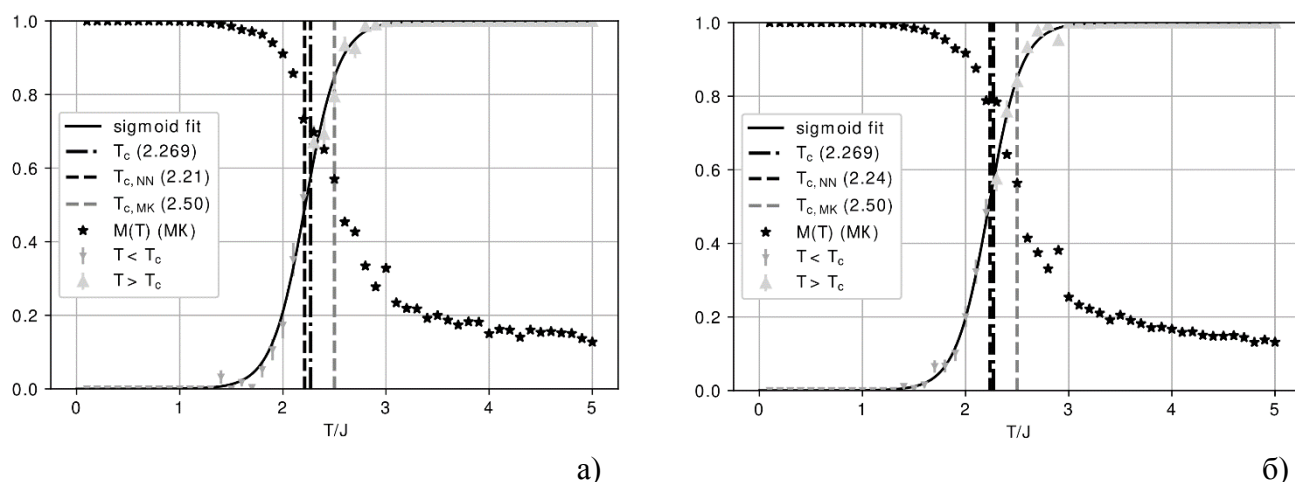


Рис. 3. Результаты расчетов T_c различными методами для системы 10×10 спинов Изинга: а) выборка равна 50, б) выборка равна 100

Как можно видеть из рис. 3 (б), при анализе большей по размеру выборки состояний, чем которая применялась для обучения, как на рис. 3 (а), были получены более точные данные, а сеть дополнительно «дообучилась» на таких данных. Уменьшение обучающей выборки также позволило сократить время численных экспериментов. В среднем можно отметить большую точность получаемых результатов по сравнению с Монте-Карло моделированием.

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, сведя их к задаче классификации спиновых состояний при разных температурах. Показана зависимость от числа Монте-Карло шагов и размера выборки для точности обучения сети и ее последующего применения, в сравнении с алгоритмом Метрополиса. Рассмотрены системы различных размеров и влияние размерного эффекта на точность получаемых результатов.

В ходе дальнейших исследований планируется применить нейронные сети к изучению более сложных моделей и решеток, в частности для изучения фрустрированных систем, таких, например, как спиновые стекла.

Список литературы

- [1] Васильев Е. В. и др. Численное моделирование двумерных магнитных скирмионных структур // Компьютерные исследования и моделирование. 2020. Т. 12. №. 5. С. 1051-1061.
- [2] Солдатов К. С. и др. Конечно-размерный скейлинг в ферромагнитных спиновых системах на решетке пирохлора // Дальневосточный математический журнал. 2020. Т. 20. №. 2. С. 255-266.

- [3] **Шаповалова К. В. и др.**, Методы канонического и мультиканонического семплирования пространства состояний векторных моделей //Дальневосточный математический журнал. 2017. Т. 17. №. 1. С. 124-130.
- [4] **Шаповалова К. В. и др.** Зачем суперкомпьютер Дальневосточному федеральному университету? // Современные наукоемкие технологии. 2017. №. 1. С. 81-87.
- [5] **Макаров А. Г. и др.** К численному расчету фрустраций в модели Изинга //Письма в Журнал экспериментальной и теоретической физики. 2019. Т. 110. №. 10. С. 700-705.
- [6] **Carrasquilla J. and Melko R.G.** Machine learning phases of matter //Nature Physics. 2017. Т. 13. №. 5. С. 431-434.
- [7] **Kenta S., et al.** Machine-Learning Studies on Spin Models //Scientific reports. 2020. Т. 10. №. 1. С. 1-6.
- [8] **Капитан В. Ю., Шевченко Ю. А., Андриященко П. Д., Нефедев К. В.** Суперкомпьютерное моделирование и численные решения задач статфизики, Изд-во ДВФУ, Вл-к. 2017. 195 с.
- [9] **Goodfellow I., Bengio Y., Courville A.** Deep learning, MIT press. 2016.
- [10] **Abadi M. et al.** Tensorflow: Large-scale machine learning on heterogeneous distributed systems //arXiv preprint arXiv:1603.04467. 2016.

References

- [1] **Vasil'ev E. V. et al.** Numerical simulation of two-dimensional magnetic skyrmion structures //Computer Research and Modeling. 2020. vol. 12. no. 5. – pp. 1051-1061.
- [2] **Soldatov K. S. et al.** Finite-size scaling in ferromagnetic spin systems on the pyrochlore lattice //Dal'nevostochnyi Matematicheskii Zhurnal [Far Eastern Mathematical Journal]. 2020. vol. 20. no. 2. pp. 255-266.
- [3] **Shapovalova K. V. et al.** Methods of canonical and multicanonical sampling of the space of vector models // Dal'nevostochnyi Matematicheskii Zhurnal [Far Eastern Mathematical Journal]. 2017. vol. 17. no. 1. pp. 124–130
- [4] **Shapovalova K. V. et al.** Why does Far East Federal University need supercomputer? // Sovremennye naukoemkie tekhnologii [Modern high technologies]. 2017. no. 1. pp. 81–87
- [5] **Makarov A. G. et al.** On the Numerical Calculation of Frustrations in the Ising Model // JETP Letters. 2019. vol. 110. no. 10. pp. 702–706.
- [6] **Carrasquilla J. and Melko R.G.** Machine learning phases of matter //Nature Physics. 2017. vol. 13. no. 5. pp. 431-434.

[7] **Kenta S., et al.**, Machine-Learning Studies on Spin Models //Scientific reports. 2020. vol. 10. no. 1. pp. 1-6.

[8] **Kapitan V.Yu., Shevchenko Yu.A., Andriushchenko P.D., Nefedev K.V.**, Supercomputer modeling and numerical solutions of statistical physics problems, Publishing house of FEFU, Vladivostok, 2017. 195 p.

[9] **Goodfellow I., Bengio Y., Courville A.** Deep learning, MIT press. 2016.

[10] **Abadi M. et al.** Tensorflow: Large-scale machine learning on heterogeneous distributed systems //arXiv preprint arXiv:1603.04467. 2016.