Extraction of properties of anisotropic spin model by deep transfer learning methods* †

D.D. Sukhoverkhova[1,2], L.N. Shchur[1,2]

[1]National Research University Higher School of Economics, 101000, Moscow, Russia
[2]Landau Institute for Theoretical Physics, 142432, Chernogolovka, Moscow region, Russia

We apply supervised deep machine learning techniques to extract properties of the anisotropic Ising model. We consider two cases of anisotropy: orthogonal and diagonal. From the predictions of the neural network, we obtained phase probability functions, from which we measured two quantities: the critical temperature and the critical exponent of the correlation length. We estimated the values of the anisotropy parameter in both cases at which the neural network predictions correctly reproduce the critical behaviour. When the anisotropy is significant, the neural network predicts phases incorrectly. We attribute this to a change in the behaviour of the correlation function. For example, in the case of diagonal anisotropy, these are oscillations of the correlation function that lead to significant deviations in the predictions.

Keywords: phase transitions, Ising model, supervised machine learning, transfer learning.

1. Introduction

Machine learning methods have been widely applied in recent decades in tasks requiring the processing of large amounts of data, including the study of statistical physics systems characterised by an extensive phase space. For example, the research in [1] focused on extracting information about the second-order phase transition in the Ising model. The authors formulated a classification problem for a neural network where the phase transition occurs between ferro- and paramagnetic phases representing two classes. By analysing the probability distribution of these phases, it is possible to estimate the critical transition temperature and some universal properties of the Ising model.

Models with the same dimensionality, symmetry, and degeneracy of the ground state form universality classes, and in the critical region models from the same class have identical properties. Using machine learning, it is possible to identify the properties of a particular universality class to which the model in question belongs [2]. It is known that the correlation length growth rate that diverges at the point of phase transition of the second kind [3], can be estimated using a neural network method. This leads to the question of the feasibility and accuracy of using such an approach to estimate the critical correlation length exponent for other models from similar universality classes. Realising this requires using the knowledge gained during training of the neural network on new data.

Traditional approaches to extract properties of spin models are Monte Carlo class methods: cluster algorithms (Wang-Landau [4], Wolff [5])and multicanonical algorithms [6]. Spin models have also been investigated by machine learning methods [1,7,8], and there are papers on transfer learning between different universality classes [9]. But to the best of authors' knowledge, no one has investigated the impact of transfer of learning in a single universality class.

We focus on transfer learning in one universality class. For this purpose, we train a convolutional neural network on samples of isotropic Ising model, and then we test the already trained network on anisotropic samples. Anisotropy in the system arises due to changes in the coupling

*This paper is the result of a research project implemented as part of the fundamental research programme of the National Research University Higher School of Economics (HSE).
†This research was supported in part through computational resources of HPC facilities at HSE University [17].

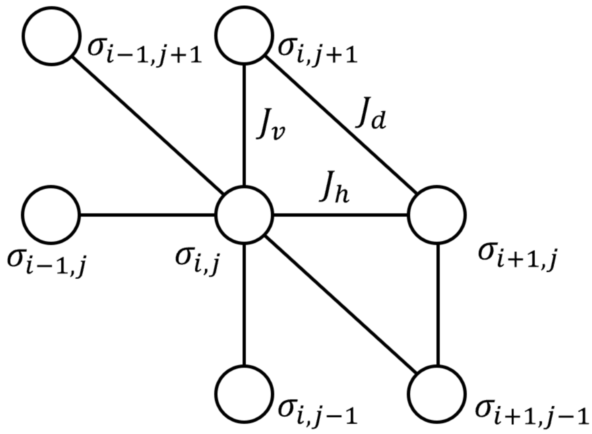


Figure 1. Spin $\sigma_{i,j}$ and its nearest neighbours of the Ising model on a triangular lattice

constants between spins. We consider two cases of anisotropic interactions: in the first case, the anisotropy parameter is the ratio of the magnitude of the vertical bond to the horizontal bond on a square lattice. In the second case, the vertical and diagonal links are equal, and the parameter is the ratio of the magnitude of the diagonal link to the horizontal link on a triangular lattice. Thus, we verify the accuracy with which universal properties of the system can be extracted using neural network analysis.

2. Methods

2.1. Models under study

For studying transfer learning applicability, we consider the Ising model on triangular lattice (Fig. 1):

$$\mathcal{H}=-\sum_{i,j=1}^{L}\sigma_{i,j}\left[J_h\sigma_{i+1,j}+J_v\sigma_{i,j+1}+J_d\sigma_{i+1,j+1}\right]. \tag{1}$$

Here the parameters J_v, J_h, J_d are the coupling constants on the vertical, horizontal and diagonal respectively, L is the system size.

When $J_v = J_h, J_d = 0$, it is the isotropic Ising model in a square lattice with the Z_4 symmetry of a lattice. Respectively, if we vary coupling constants, symmetry is broken and the system becomes anisotropic.

We consider two cases of anisotropy, in each of which the spatial Z_2 symmetry is preserved:

- $J_v = J_h = J = 1,\ \kappa_1 = J_d/J \neq 0$ – diagonal anisotropy
- $J_d = 0,\ \kappa_2 = J_v/J_h \neq 1$ – orthogonal anisotropy

Each instanton of the Ising model exists at its own temperature T, depending on which the system can be in an ordered (ferromagnetic) or disordered (paramagnetic) state. Orderedness is characterised by spontaneous magnetisation in the absence of an external magnetic field.

2.2. Neural network output

We feed the instantaneous samples (snapshots) of the model 1 to a convolutional neural network (see 'Neural network pipeline' section) and train the network to solve the classification problem.

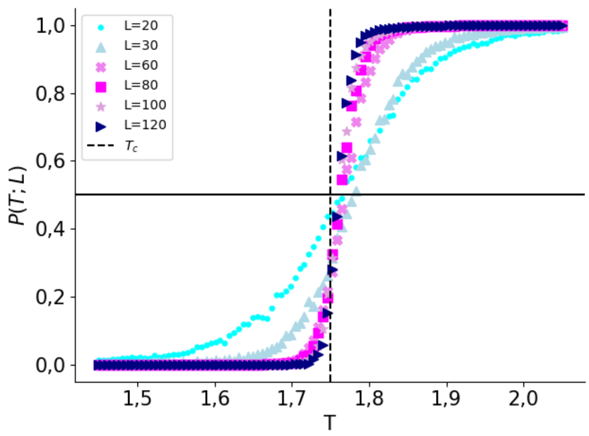


Figure 2. Probability function of samples belonging to paramagnetic phase $P(T;L)$; $\kappa_1 = -0.3$

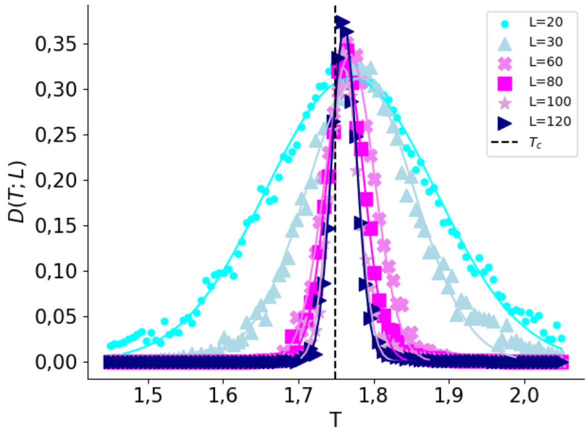


Figure 3. Standard deviations of the probability of samples belonging to paramagnetic phase $D(T;L)$; $\kappa_1 = -0.3$

In the Ising model, a phase transition of the 2nd kind between the ferromagnetic and paramagnetic phases is observed at the critical point T_c, also known as the phase transition temperature. The phase transition temperature of this model is known from the analytical solution [10]:

$$\sinh\frac{2J_v}{k_BT_c}\ \sinh\frac{2J_h}{k_BT_c}+\sinh\frac{2J_h}{k_BT_c}\sinh\frac{2J_d}{k_BT_c}+\sinh\frac{2J_d}{k_BT_c}\sinh\frac{2J_v}{k_BT_c}=1, \tag{2}$$

$$J_v+J_h>0,\ \ J_h+J_d>0,\ \ J_d+J_v>0.$$

Knowing $T_c(J_v, J_h, J_d)$, we divide all data into two classes: 0 – ferromagnetic ($T < T_c$), 1 – paramagnetic ($T > T_c$). After that, we train the neural network to separate the incoming data into two classes.

The output of the neural network is $p_i(T;L)$ – the probability of a sample i with temperature T and size $L \times L$ to be in the paramagnetic phase. Having obtained the predictions for N samples, we construct the paramagnetic phase probability functions $P(T;L)$ 3 and the second moments of this function $D(T;L)$ 4:

$$P(T;L)=\frac{1}{N}\sum_{i=1}^{N}p_i(T;L)\ , \tag{3}$$

$$D(T;L)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i(T;L))^2-\left(\frac{1}{N}\sum_{i=1}^{N}p_i(T;L)\right)^2}\ . \tag{4}$$

2.3. Temperature estimation

We build functions $P(T;L)$ and $D(T;L)$ for several lattice sizes and over a given temperature range estimate the phase transition temperature T_c° found by the neural network: the distribution $D(T;L)$ at each L is approximated by a Gaussian function with mean μ and standard deviation σ. We consider $T_c^- = \mu(L)$ as an estimate of the critical temperature.

All estimates are generated by testing the neural network on finite grid dimensions L. We then make an estimate of the corresponding temperature in the thermodynamic limit based on the shift of the critical temperature in finite dimensions $T_c^-(\infty) = T_c^-(L) + a/L$.

2.4. Critical exponent estimation

It is known [3] that in the universality class of the two-dimensional Ising model, the correlation length ξ diverges at the phase transition point: $\xi \propto \tau^{-1/\nu}$ with critical exponent $\nu = 1$, and $\tau = (T - T_c)/T_c$ is the reduced temperature.

We find an estimate of the exponent ν using the hypothesis [2] that the width of $\sigma(L)$ at finite lattice sizes behaves in the same way as the width of thermodynamic functions [11, 12]:

$$\sigma(L) \propto bL^{-1/\nu}.$$

3. Neural network pipeline

The data are generated using the Metropolis algorithm, with each time point being a black and white image or snapshot. The thermalization time is $20 \times L^{2.15}$ [13]. Once equilibrium is reached, each snapshot is saved in every $2 \times L^{2.15}$ Monte Carlo step, corresponding to $L \times L$ local spin flips.

Each data set is created in the range $T_c \pm 0.3$, comprising 100 temperature points uniformly distributed in increments of 6×10^{-3}. The value of T_c varies with J_v according to the formula 2. For each temperature T in the specified range, we store $N = 2048$ snapshots under isotropic sampling, where $J_v = J_h = 1$. In the case of generating anisotropic datasets when $J_v \neq J_h$, we only store $N = 512$ snapshots. As a result, each dataset contains $N \times 100$ images.

We use a convolutional neural network (CNN) architecture, which includes one convolutional layer, two fully connected layers, and ReLU activation between them [2]. We train neural networks within one epoch to avoid overtraining [14].

The neural networks are trained on isotropic samples with parameters $J_d = 0$, $J_v = J_h = 1$. The pre-trained networks are then tested on each test sample: in the case of diagonal anisotropy, we vary the parameter κ_1 from -0.7 to 1.0 inclusive in steps of 0.1; $J_v = J_h = 1$; in the case of orthogonal anisotropy we change the parameter κ_2: 1, 3/4, 1/2, 1/8, 1/16; $J_d = 0$.

4. Results

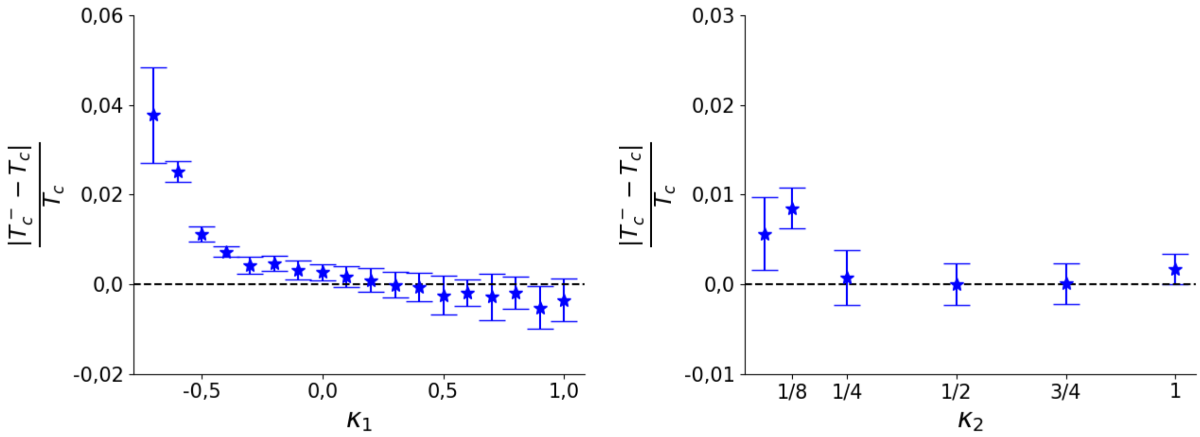
As described above, we estimated the phase transition point T_c^- and the critical exponent $1/\nu$ for different anisotropy parameters κ_1 and κ_2. The results of the critical temperature estimation are shown in the graphs 4, the $1/\nu$ estimates are collected in the Tables 1–3 and depicted in the graphs 5.

In the results of estimation of both quantities there is a systematic deviation from theoretical values: growth of the relative error of critical temperature and deviation of critical exponent $1/\nu$ from the theoretical value of 1. Such an effect is observed when: $\kappa_1 \leq -0.5$; $\kappa_2 < 1/4$.

Table 1. Estimates of the critical exponent of correlation length ν obtained from analysing the width $\sigma(L)$ of the function $D(T; L)$: the second line is full width, the right half-width $\sigma_r(L)$ of this function is the third line, the left half-width $\sigma_l(L)$ is the fourth line; $\kappa_1 \in [-0.7, 0.0]$

κ_1	**–0.7**	**–0.6**	**–0.5**	**–0.4**	**–0.3**	**–0.2**	**–0.1**	**0.0**
$\frac{1}{\nu}$	0.52(15)	0.68(9)	0.88(10)	1.02(5)	1.06(2)	1.07(4)	1.11(2)	1.09(2)

On a triangular lattice with diagonal anisotropy, we observe a deviation of the critical temperature and correlation length exponent when $\kappa_1 < -0.4$. For example, at $-0.4 < \kappa_1 \leq 1.0$ the temperature estimates $\mu(L)$ lie on a curve corresponding to the normalised critical temperature known analytically [15] (Fig. 6a). As the parameter κ_1 decreases, the critical temperature decreases accordingly and the phase diagram shifts towards the region containing the disorder



a) b)

Figure 4. Relative error of the critical temperature estimate T_c^-: a) $\kappa_1 \in [-0.7, 1.0]$, b)$\kappa_2 = 1/16, ..., 1$

Table 2. Estimates of the critical exponent of correlation length ν obtained from analysing the width $\sigma(L)$ of the function $D(T; L)$: the second line is full width, the right half-width $\sigma_r(L)$ of this function is the third line, the left half-width $\sigma_l(L)$ is the fourth line; $\kappa_1 \in [0.1, 1.0]$

κ_1	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
$\frac{1}{\nu}$	1.11(5)	1.11(5)	1.10(5)	1.08(3)	1.11(5)	1.07(5)	1.06(5)	1.09(3)	1.09(4)	1.08(7)

Table 3. Estimates of the critical exponent of correlation length ν obtained from analysing the width $\sigma(L)$ of the function $D(T; L)$: the second line is full width, the right half-width $\sigma_r(L)$ of this function is the third line, the left half-width $\sigma_l(L)$ is the fourth line; $\kappa_2 = 1/16, ..., 1.0$

κ_2	**1**	**3/4**	**1/2**	**1/4**	**1/8**	**1/16**
$\frac{1}{\nu}$	1.10(3)	1.07(4)	1.03(3)	1.06(4)	0.87(7)	0.66(11)

temperature. In this region, pairwise correlated spin oscillations are observed in the system, which arise due to the diagonal antiferromagnetic interaction. During training, the neural network was fed with data that did not contain such oscillations, so the results of the estimation of the quantities of interest deviate from the theoretical values in the same regions.

As for the orthogonal anisotropy, we assume that the deviations at small parameters κ_2 are due to the peculiarities of the correlation function in this region. In [16], a decomposition of the correlation function $< \sigma_{00}\sigma_{ij} >$ over radially measured distance R was obtained (sign + or $-$ stands for temperature region: greater or less than T_c):

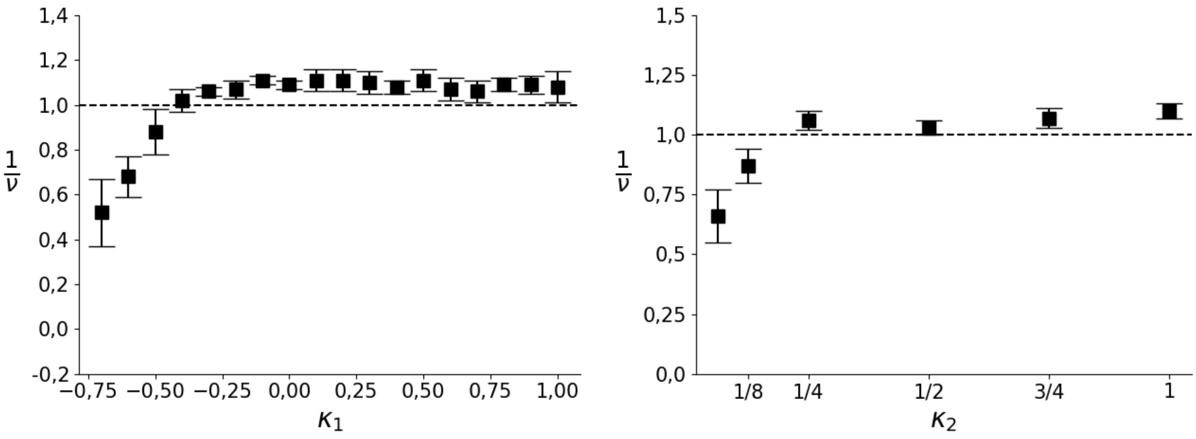
$$< \sigma_{00}\sigma_{ij} >= F_{\pm}(t)/R^{1/4} + F_{1\pm}(t)/R^{5/4} + o(R^{-5/4}). \tag{5}$$

Here R is the radial radius, in the isotropic case $R = L\sqrt{2}$; t is the reduced temperature in variables $z1 = \tanh \beta J_h,\ z2 = \tanh \beta J_v$:

$$t = |z_1 z_2 + z_1 + z_2 - 1| \left[z_1 z_2 (1 - z_1^2)(1 - z_2^2) \right]^{1/4} R. \tag{6}$$

For example, we construct the ratio R_1:

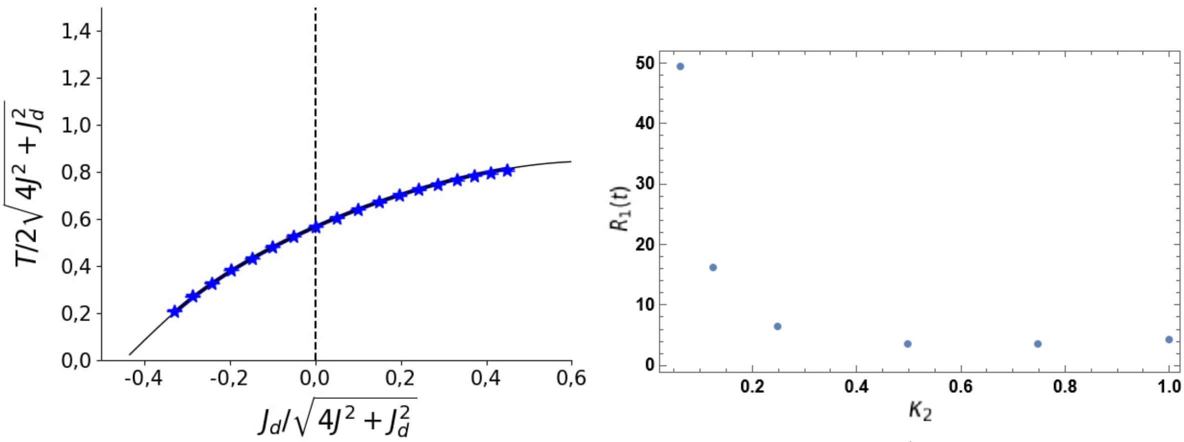
$$R_1 = \frac{F_{1-}(t)}{tF_{-}(t)}. \tag{7}$$



a) b)

Figure 5. Estimation of the critical correlation length exponent $(1/\nu)_-$: black squares, $1/\nu$, full width; red triangles, $(1/\nu)_r$, right half-width; green stars, $(1/\nu)_l$, left half-width; a) $\kappa_1 \in [-0.7, 1.0]$, b) $\kappa_2 = 1/16, ..., 1$

Fig. 6b repeats Fig. 2 from the paper [16], and at the points $\kappa_2 = 1/8, 1/16$, an increase in the R_1 ratio is noticeable. This means that in samples at lower temperatures than T_c possessing orthogonal anisotropy, the influence of the correction term $F_{1-}(t)R^{-5/4}$ increases, which affects the spatial behaviour of the correlation function, which is also not taken into account when training the neural network.



a) b)

Figure 6. a) Estimations of critical temperature T_c^- on the same plot with precise critical temperature from paper [15]; b) ratio 7 vs anisotropy parameter κ_2

5. Discussion

In this work, we investigated the applicability of transfer learning within one universality class of the two-dimensional Ising model. We trained a convolutional neural network to discriminate between the two phases of an isotropic model, and then tested the pre-trained network on anisotropic samples. From the predictions of the network, we extracted estimates of the critical temperature and the critical correlation length exponent. The results showed that we estimate both quantities according to their theoretical values (within the statistical error) when transferring the training, but the region of correctness is limited.

The hypothesis is that at strong anisotropy there are non-trivial effects related to spatial oscillations occurring in the system. In the case of diagonal anisotropy, antiferromagnetic diagonal interactions are influential, leading to frustrations due to the unattainability of the ground state. In the case of orthogonal anisotropy, the increase of correction term in the correlation function leads to a different behaviour on the short scales, limiting the applicability of transfer learning. Further work on this problem will consist in testing the hypotheses put forward.

References

1. Carrasquilla J., Melko R.G. Machine learning phases of matter // Nature Physics. 2017. Vol. 13, no. 5. P. 431–434. DOI: 10.1038/nphys4035.
2. Chertenkov V., Burovski E., Shchur L. Finite-size analysis in neural network classification of critical phenomena // Physical Review E. 2023. Vol. 108, no. 3. P. L032102. DOI: 10.1103/PhysRevE.108.L032102.
3. Onsager L. Crystal statistics. I. A two-dimensional model with an order-disorder transition // Physical Review. 1944. Vol. 65, no. 3-4. P. 117. DOI: 10.1103/PhysRev.65.117.
4. Wang F., Landau D.P. Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram // Phys. Rev E. 2001. Vol. 64. P. 056101. DOI: 10.1103/PhysRevE.64.056101.
5. Wolff U. Collective Monte Carlo updating for spin systems // Phys. Rev. Lett. 1989. Vol. 62, no. 4. P. 361–364. DOI: 10.1103/PhysRevLett.62.361.
6. Berg B.A., Neuhaus T. Multicanonical ensemble: A new approach to simulate first-order phase transitions // Physical Review Letters. 1992. Vol. 68, no. 1. P. 9. DOI: 10.1103/PhysRevLett.68.9.
7. van Nieuwenburg E.P.L., Liu Y.H., Huber S.D. Learning phase transitions by confusion // Nature Physics. 2017. Vol. 13, no. 5. P. 435–439. DOI: 10.1038/nphys4037.
8. Ng K.K., Yang M.F. Unsupervised learning of phase transitions via modified anomaly detection with autoencoders // Physical Review B. 2023. Vol. 108, no. 21. P. 214428. DOI: 10.1103/PhysRevB.108.214428.
9. Chertenkov V., Shchur L. Supervised and Transfer Learning for Phase Transition Research // Russian Supercomputing Days. Springer Nature Switzerland, Cham, 2024. P. 434–449. DOI: 10.1007/978-3-031-78459-0_31.
10. Houtappel R.M.F. Order-disorder in hexagonal lattices // Physica. 1950. Vol 16, no. 5. P. 425–455. DOI: 10.1016/0031-8914(50)90130-3.
11. Fisher M.E., Ferdinand A.E. Interfacial, boundary, and size effects at critical points // Physical Review Letters. 1967. Vol. 19, no. 4. P. 169. DOI: 10.1103/PhysRevLett.19.169.
12. Ferdinand A.E., Fisher M.E. Bounded and inhomogeneous Ising models. I. Specific-heat anomaly of a finite lattice // Physical Review. 1969. Vol. 185, no. 2. P. 832. DOI: 10.1103/PhysRev.185.832.
13. Chertenkov V., Shchur L. Universality classes and machine learning // Journal of Physics: Conference Series. 2021. Vol. 1740, no. 1. P. 012003. DOI: 10.1088/1742-6596/1740/1/012003.
14. Chertenkov V.I. PhD thesis, National Research University Higher School of Economics (2024).

15. Stephenson J. Ising model with antiferromagnetic next-nearest-neighbor coupling: Spin correlations and disorder points // Physical Review B. 1970. Vol. 1, no. 11. P. 4405. DOI: 10.1103/PhysRevB.1.4405.

16. Wu T.T. et al. Spin-spin correlation functions for the two-dimensional Ising model: Exact theory in the scaling region // Physical Review B. 1976. Vol. 13, no. 1. P. 316. DOI: 10.1103/PhysRevB.13.316.

17. Kostenetskiy P.S., Chulkevich R.A., Kozyrev V.I. HPC Resources of the Higher School of Economics // Journal of Physics: Conference Series. 2021. Vol. 1740, no. 1. P. 012050. DOI: 10.1088/1742-6596/1740/1/012050.