Machine Learning Domain Adaptation in Spin Models with Continuous Phase Transitions

Vladislav Chertenkov[1,2] and Lev Shchur[1,2]
[1] *Landau Institute for Theoretical Physics, 142432 Chernogolovka, Russia and*
[2] *HSE University, 101000 Moscow, Russia*

The main question raised in the letter is the applicability of a neural network trained on a spin lattice model in one universality class to test a model in another universality class. The quantities of interest are the critical phase transition temperature and the correlation length exponent. In other words, the question of transfer learning is how "universal" the trained network is and under what conditions. The traditional approach with training and testing spin distributions turns out to be inapplicable for this purpose. Instead, we propose to use training and testing on binding energy distributions, which leads to successful estimates of the critical temperature and correlation length exponent for cross-tested Baxter-Wu and Ising models belonging to different universality classes.

Introduction. — Machine learning [1] is a promising approach with possible applications in statistical physics. It has already been established that supervised learning with classification of the spin distribution into ferromagnetic (FM) or paramagnetic (PM) phases can yield estimates of the critical temperature [2] and correlation length exponent [3] when training and testing the same model. A natural question is: can we train a neural network (NN) with a model in one universality class and use it to estimate the critical temperature and correlation length exponent of a model of another universality class? With the right choice of the data set, we answered this question in the affirmative.

We use a recently developed finite-size scaling approach [3] to analyze *variation* $V(T)$ of the NN output function $P(T)$ - it has been shown that extracting the correlation length and critical temperature from $V(T)$ is very robust both when applied to different physical models and when using different neural network (NN) architectures, as long as training and testing are performed on the same model.

To the letter, we show that the problem of learning/testing spin distributions involves two aspects. Firstly, there are the physical properties of ordered states, which lie in the degeneracy of the ground state. For the Ising model, the degeneracy is two, and the system below T_c can equally have most spins in the +1 or -1 state. In the case of the Baxter-Wu model, four combinations of spins on triangular plaques give the same ground state energy, and this is a fourfold degeneracy. The ground state degeneracy separates systems with the same dimensionality and the same number of order parameter components into different universality classes [4]. In practice, this means a different set of values of the critical exponents.

Second, systems in the same universality class can differ in the symmetry of the interaction between spins, which is reflected in the peculiarities of the local configurations of spins. For example, the universality class of the 4-component Potts model [5] includes the Baxter-Wu model [6], special cases of the Ashkin-Teller model [7], and the Turban model [8]. All four models have the same set of critical exponents and indistinguishable properties of thermodynamic functions at temperatures in the neighborhood of the critical temperature [9]. From the results of applying machine learning to second-order phase transitions [2, 3] , we can conclude that the neural network somehow senses correlations in the system, and it is not clear how NNs will respond to local features of the models.

Our problem is the part of the very broad problem of the transfer learning [10], and generally it is connected with the problem of domain adaptation, which arises when there are several related domains containing a shift associated with differences in feature space distributions. Snapshots of spin distributions for different models at temperatures below the critical temperature tend to be very different from each other, and NN cannot correctly predict the probability that specific spin distributions belong to the ferromagnetic phase. The goal is to train a robust model to generalize the common properties of the domains, such as phase transition point and set of the critical exponents.

We have analyzed several possible representations of the datasets used for the training/testing procedure, including traditional spins snapshots [2], and found that they do not provide acceptable estimates of the critical temperature and correlation length exponent during cross-domain transfer learning (see details in the Supplemental Materials). Fortunately, converting spin snapshots into energy snapshots, which are then used in the training/testing process, leads to reasonably good estimates of the critical temperature and correlation length exponent during cross-domain testing of the Baxter-Wu model using the Ising-trained neural network. Conversely, training on the Baxter-Wu dataset and testing on the Ising dataset also yields good results.

Statistical physics models.— We focus on the two two-dimensional models, the Ising model [11] and the Baxter-Wu model [6]. The following representation of the Hamiltonians is the key to the chosen representation with energy couplings in the datasets used for training and testing. Only ferromagnetic couplings $J>0$ are considered and periodic boundary conditions employed for all models. In the following text, we set the coupling constant J equal to one, measuring the energy in units of J.

Two-dimensional Ising (IS) model on a square lattice defined by the Hamiltonian

$$\mathcal{H}_{ising} = -J \sum_{(i,j)} \left[\sigma_{i,j}\sigma_{i+1,j} + \sigma_{i,j}\sigma_{i,j+1}\right]. \quad (1)$$

Two-dimensional Baxter-Wu (BW) model on a triangular lattice defined by the Hamiltonian

$$\mathcal{H}_{bw} = -J \sum_{(i,j)} \left[\sigma_{i,j}\sigma_{i+1,j}\sigma_{i+1,j+1} + \sigma_{i,j}\sigma_{i,j+1}\sigma_{i+1,j+1}\right], \quad (2)$$

and the lattice can be easily understood as a square lattice with a right-to-down diagonal, and the first term in brackets is the product of the spins of the up triangle and the second term of the down triangle.

The Baxter-Wu model belongs to the universality class of the four-component Potts model. The Baxter-Wu model is chosen because it lacks [12] multiplicative logarithmic corrections [13, 14], which are quite pronounced in the four-component Potts model and spoil its critical behaviour. The Ising model belongs to a different class of universatility. Deep machine learning allows us to estimate the critical temperature T_c [2] and the value of the correlation length exponent ν [3], which are known analytically and equal to 1 for the Ising model [11] and 2/3 for BW [6].

Spin snapshot datasets — The datasets are generated with conventional Markov chain Monte Carlo methods and consist of spin configurations on the lattice [15], the so-called spin snapshots. Each spin snapshot is taken after the previous one after a time larger than the correlation time, making them statistically independent (details can be found in the article [3]). The snapshots are used to train and test NNs. Mathematically, a spin snapshot is a $L\times L$ matrix, where L is the linear size of the system. The elements of the matrix are the values of spins. In the case of Ising and BW model these are the values 1 and -1.

Supervised learning. — We performed transfer learning between these models in all possible combinations of training and testing, using two neural network architectures: convolutional neural network (CNN) [16] and deep convolutional residual network (ResNet-10) [17] with 10 layers.

We first trained the networks with 1000 snapshots for each model and each parameter set $(T;L)$, marking them as FM if $T < T_C$ and PM otherwise. The trained neural networks are used to test models with N=500 snapshots modeled for the same parameter set. The output of the FM phase prediction neuron is denoted by $f_i(T;L)$, which has a value between 0 and 1, where 1 means that NN recognized the snapshot indexed i, $(i=1,2,\ldots,N)$ as definitely belonging to the FM phase, and 0 if it does not definitely belong to the FM phase. The values $f_i(T;L)$ are used to estimate corresponding probability [2]

$$P(T;L) = \frac{1}{N_t}\sum_{i=1}^{N_t} f_i(T;L) \quad (3)$$

and the variation (VOT) [3]

$$V(T;L) = \frac{1}{N_t}\sum_{i=1}^{N_t} \left(f_i(T;L)\right)^2 - P(T;L)^2. \quad (4)$$

Test results using spin data sets. — To analyse the resulting functions, we use the procedure proposed in [3], which uses a Gaussian fit of the VOT function and analyses the mean $\mu(L)$ and width $\sigma(L)$ as functions of lattice size. It is based on analogy with the known facts [18, 19] from statistical mechanics that in the critical region there is a shift in the pseudocritical temperature $\mu(L)-T_c \propto 1/L^{1/\nu}$, and the typical width behaves as $\sigma(L) \propto 1/L^{1/\nu}$.

The results of testing the Ising model and the Baxter-Wu model with the NN trained on *same* model are shown in the table I, in the second and fourth columns. The third and fifth columns are the difference between the estimated critical temperature and the exact critical temperature $\Delta=|T^{*}-T_c|$ divided by the statistical error ϵ of the linear fitting [20]. It can be seen that IS@BW transfer learning predicts the critical temperature with an accuracy approximately equal to that of direct learning, and the abbreviation stands for testing the Ising model on the NN trained by the Baxter-Wu model. We do not provide estimates for testing/training BW@IS because fitting the VOT function to a Gaussian is not possible. Estimation of the correlation length exponent is not possible for both test/training cases, IS@BW and BW@IS. The upper panel of the figure 1 shows the scatter of the ferromagnetic phase prediction in the Baxter-Wu model, which is due to the difference in the degeneracy of the ground state of the tested model and the Ising model whose snapshots were used for training.

NN	T^{*}, IS	Δ/ϵ	T^{*}, BW	Δ/ϵ	T^{*}, IS@BW	Δ/ϵ
CNN	2.273(6)	0.7	2.2687(2)	2.0	2.280(6)	1.8
ResNet-10	2.267(2)	1.2	2.2690(2)	1.1	2.301(29)	1.1

TABLE I. Estimation of the critical temperature using *spin configurations* by testing the Ising model (IS) and the Baxter-Wu (BW) model with a NN trained on the *same* model (second and fourth columns) and testing the Ising model using a NN trained on *another* model, Baxter-Wu (last two columns). For the meaning of Δ and ϵ, see text.

Energy snapshot datasets — Due to the failure of cross-domain testing using spin snapshots, we propose to use a different representation of the input datasets, the energy-based representation. The energy-based representation reflects spin interactions in the Hamiltonians (1)-(2).

For the Ising model datasets, using spin snapshots, we form two $L \times L$ matrices with elements equal to the horizontal coupling energy with elements $e^1_{i,j} = -\sigma_{i,j}\sigma_{i+1,j}$ in the first matrix and with elements equal to the vertical coupling energy with elements $e^2_{i,j} = -\sigma_{i,j}\sigma_{i,j+1}$ in the second matrix.

For BW model datasets, the first matrix elements calculated with the first term in Expr. (2) $e^1_{i,j} = -\sigma_{i,j}\sigma_{i+1,j}\sigma_{i+1,j+1}$ and the second matrix elements calculated with the second term $e^2_{i,j} = -\sigma_{i,j}\sigma_{i,j+1}\sigma_{i+1,j+1}$.

These representations result in matrix elements equal to -1 or 1 for all models. Our energy snapshot approach is an advance on the problem of domain adaptation of different models, including transfer learning between universality classes. It is well suited for cross-domain training between the Ising and the Baxter-Wu models.

Energy datasets results. — The results of critical temperature estimation when testing the Ising model and the Baxter-Wu model with a network trained on *same* model and using the energy data sets is presented in the Table II. The results should be compared with the data in table I, and the bias in the fit is smaller for vigorous datasets, although we should not take this characteristic too seriously, as we have not yet found a way to estimate the systematic error of the approach.

In contrast to cross-domain testing with the *spin* datasets, the cross-domain testing with the *energy* datasets yields satisfactory estimate of the critical temperature for both IS@BW and BW@IS combinations, which is presented in table III and should be compared with the results in the last two columns of the table I. In addition, such approach makes it possible to extract estimates for the correlation length exponent ν.

NN	T^*, IS	Δ/ϵ	T^*, BW	Δ/ϵ
CNN	2.266(8)	0.4	2.2696(7)	0.6
ResNet-10	2.263(8)	0.8	2.268(1)	1

TABLE II. Same as in the table I for training and testing the *same* model but using *energy* datasets.

NN	T^*, IS@BW	Δ/ϵ	T^*, BW@IS	Δ/ϵ
CNN	2.226(24)	1.8	2.2694(2)	1
ResNet-10	2.214(21)	2.6	2.2686(5)	1.2

TABLE III. Same as in the table I for training and testing the *another* model but using *energy* datasets.

From the width of the VOT function, expr. (4), we obtain an estimate of the exponent of the correlation length using three methods: fitting the entire dataset $V(T;L)$ as a function of temperature T using a Gaussian and estimating $\sigma(L)$, separately fitting the left wing of the data gives an estimate of $\sigma^-(L)$, and separately fitting the right wing gives an estimate of $\sigma^+(L)$. The result of the work [3] states that the width and half-width scale with an exponent close to the inverse exponent of the correlation length ν. Tables IV and V show the $1/\nu$ estimates for two models – Ising and Baxter-Wu. It can be seen that the estimates within the errors are consistent with those obtained from spin snapshots for in-domain training/testing and presented in Tables 1 and 3 of the article [3].

Fig. 1 shows how the energy dataset representation as opposed to the spin dataset, enables cross-domain transfer learning of the phase transition for BW@IS combination of models. While the behaviour of the function $V(T;L)$ calculated using the spin domain classification does not provide a way to analyse critical behaviour of the Baxter-Wu model using the NN trained on the Ising model, the use of the energy domain gives clear predictions and the possibility to analyse the critical behaviour with acceptable accuracy.

Thus, we obtain for the first time that an NN trained on a dataset of one universality class can be used to predict the correlation length values of a model in another universality class. And this is possible by using an appropriate representation of the dataset. In our case, these are datasets with local energy distribution.

IS@IS	$1/\nu_\sigma$	$1/\nu_{\sigma^-}$	$1/\nu_{\sigma^+}$
CNN	1.12(3)	1.15(5)	1.07(2)
ResNet-10	1.15(3)	1.17(6)	1.09(11)
IS@BW	$1/\nu_\sigma$	$1/\nu_{\sigma^-}$	$1/\nu_{\sigma^+}$
CNN	1.16(5)	1.31(8)	0.98(1)
ResNet-10	0.85(6)	1.13(11)	0.78(15)

TABLE IV. Estimates of the inverse correlation length exponent $1/\nu$ for testing the Ising model on NN trained on the Ising model IS@IS and on NN trained on the Baxter-Wu model IS@BW. Energy datasets used.

Discussion.–

The main conclusion that can be drawn from our study is that choosing an appropriate domain to represent data for the purpose of knowledge transfer in machine learning is not straightforward and obviously predictable. We had to go through many options to select a suitable domain and possible pairs of statistical physics models for the knowledge transfer process. So far, we have been able to find one pair and one data representation domain to successfully demonstrate knowledge transfer between models from different universality classes. Nevertheless,

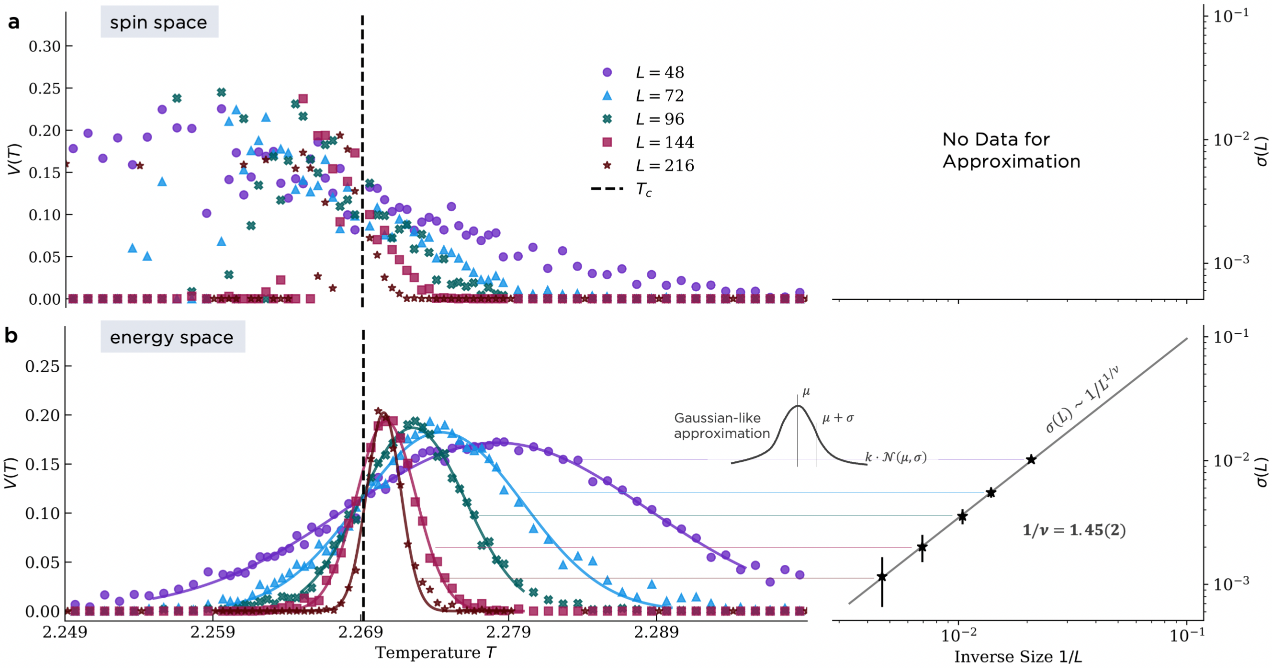

FIG. 1. Cross-domain transfer learning between Ising and Baxter-Wu models using CNNs. **a)** The variation $V(T)$ of the FM phase prediction for the Baxter-Wu model trained on the Ising model (BW@IS) using the *spin domain* cannot be approximated. **b)** The variation $V(T)$ of the same testing/training combination of models BW@IS on the *energy domain* with Gaussian approximation and estimation of the inverse correlation length exponent $1/\nu$ from $\sigma(L)$ scaling.

BW@BW	$1/\nu_\sigma$	$1/\nu_{\sigma-}$	$1/\nu_{\sigma+}$
CNN	1.48(5)	1.61(10)	1.52(3)
ResNet-10	1.50(7)	1.62(17)	1.52(4)
BW@IS	$1/\nu_\sigma$	$1/\nu_{\sigma-}$	$1/\nu_{\sigma+}$
CNN	1.45(2)	1.51(2)	1.42(6)
ResNet-10	1.45(3)	1.45(10)	1.47(4)

TABLE V. Estimates of the inverse correlation length exponent $1/\nu$ for testing the Baxter-Wu model on NN trained on the Baxter-Wu model BW@BW and on NN trained on the Ising model BW@IS. Energy datasets used.

there is one such example, and it led to transfer learning between two models in two universality classes, with a satisfactory transition temperature estimate and an estimate of the critical exponent of the correlation length.

It should be noted that many groups have previously reported critical temperature estimates using the spin representation of various models [2, 3, 21–27]. This is not surprising since they used a binary classification procedure that is very sensitive to the number of training epochs, and starting from a certain number of epochs the neural network is almost perfectly trained to classify the snapshots into ordered and disordered phases, leading to a good estimate of the transition temperature. A more detailed discussion is given in [28]. An estimate of the correlation length cannot be extracted from the NN outputs with sufficient accuracy - the network only solves the problem of classifying domains into two phases.

Research supported by the Russian Science Foundation grant 22-11-00259.

The simulation was carried out using the computing resources of high-performance computing at the National Research University Higher School of Economics.

[1] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Reviews of Modern Physics **91**, 045002 (2019).

[2] J. Carrasquilla and R.G. Melko, Machine learning phases of matter, Nat. Phys. **13**, 431 (2017).

[3] V. Chertenkov, E. Burovski, and L. Shchur, Finite-size analysis in neural network classification of critical phenomena, Phys. Rev. E **108**, L031102 (2023).

[4] V. Privman, P.C. Hohenberg, and A. Aharony, in *C. Domb, J. L. Lebowitz (Eds.), Phase Transitions and Critical Phenomena,* vol. 14, Academic Press, New York (1991).

[5] R. B. Potts, Some generalized order-disorder transformations, Cambridge Philos. Soc. **48**, 106 (1952).

[6] R. J. Baxter and F. Y. Wu, Ising model on a triangular lattice with three-spin interactions. I. The eigenvalue equation, Australian Journal of of Physics **27**, 357 (1974).

[7] J. Ashkin and E. Teller, Statistics of two-dimensional lattices with four components, Phys. Rev. **64**, 178 (1943).
[8] L. Turban, Self-dual anisotropic two-dimensional Ising models with multispin interactions, J. de Physique Lettres **43**, 259 (1982).
[9] V. Chertenkov and L. Shchur, Universality classes and machine learning, J. Phys.: Conf. Ser. **1740**, 012003 (2021).
[10] S. Pan and Q. Yang, A survey on transfer learning, IEEE Transactions on knowledge and data engineering **22**, 1345 (2009).
[11] L. Onsager, Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition, Phys. Rev. **65**, 117 (1941).
[12] L. N. Shchur and W. Janke, Critical amplitude ratios of the Baxter–Wu model, Nucl. Phys. B **840**, 491 (2010).
[13] J. L. Cardy, M. Nauenberg, and D. J. Scalapino, Scaling theory of the Potts-model multicritical point, Phys. Rev. B **22**, 2560 (1980).
[14] L. N. Shchur, B. Berche, and P. Butera, Numerical revision of the universal amplitude ratios for the two-dimensional 4-state Potts model, Nucl. Phys. B **811**, 491 (2009).
[15] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, England, 2014).
[16] K. O'Shea and R. Nash, An introduction to convolutional neural networks, arXiv:1511.08458.
[17] K. He, X. Zhang, S. Ren, and J. Sun, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR (Las Vegas, 2016), p. 770.
[18] M. E. Fisher and A. E Ferdinand, Interfacial, boundary and size effects at critical points, Phys. Rev. Lett. **19**, 169 (1967).
[19] A. E. Ferdinand and M. E. Fisher, Bounded and Inhomogeneous Ising Models. I. Specific-Heat Anomaly of a Finite Lattice, Phys. Rev. B **185**, 832 (1969)
[20] The data obtained are in agreement with previously published [3], and the differences in the figures are explained by different data sets used in the analysis and different lattice sizes, which in this case are calculated for linear sizes $L = 48, 72, 96, 144, 216$, taking into account the symmetry of the Baxter-Wu model [12].
[21] J. Zhang, B. Zhang, J. Xu, W. Zhang, and Y. Deng, Machine learning for percolation utilizing auxiliary Ising variables, Phys. Rev. E **105**, 024144 (2022).
[22] C. Alexandrou, A. Athenodorou, C. Chrysostomou, and S. Paul, The critical temperature of the 2D-Ising model through deep learning autoencoders, The European Physical Journal B **93**, 1 (2020).
[23] A. Morningstar and R.G. Melko, Deep learning the Ising model near criticality. J. Mach. Learn. Res. **18**, 1 (2018).
[24] S. Efthymiou, M. J. Beach, and R. G. Melko, Super-resolving the Ising model with convolutional neural networks, Physical Review B, **99**, 075113 (2019).
[25] W. Hu, R. R. Singh, R. T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination, Phys. Rev. E **95**, 062122 (2017)
[26] A. Canabarro, F. F. Fanchini, A. L. Malvezzi, R. Pereira, R. and Chaves, Unveiling phase transitions with machine learning, Physical Review B, **100**, 045129 (2019).
[27] K. Shiina, H. Mori, Y. Okabe, and H. K. Lee, Machine-learning studies on spin models, Scientific reports **10**, 2177 (2020).
[28] D. Sukhoverkhova, V. Chertenkov, E. Burovski, and L. Shchur, Validity and Limitations of Supervised Learning for Phase Transition Research, LNCS **14389**, 314 (2023).