Deep machine learning investigation of phase transitions

Vladislav Chertenkov [1,2]
Graduate student

Evgeni Burovski [1,2]
Associate Professor

Lev Schur [1,2]
Department Head, Professor

[1] Landau Institute for Theoretical Physics, Chernogolovka, Russia
[2] HSE University, Moscow, Russia

Program

1	BRIEF INTRODUCTION	We present two spin models and describe some details of the data generation.	9 MIN
2	MACHINE LEARNING	We describe the deep learning approach we use for the analysis.	3 MIN
3	RESULTS & OUTLOOK	We present the results of our investigation and discuss the prospects for further research.	4 MIN
4	QUESTIONS SECTION	We answer your questions.	4 MIN

Spin models

Spin models

spin up $\sigma = +1$

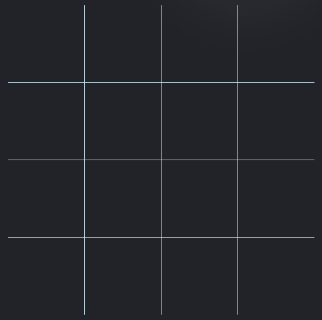
spin down $\sigma = -1$

Spin models

spin up $\sigma = +1$

spin down $\sigma = -1$

Lattice
size = L x L

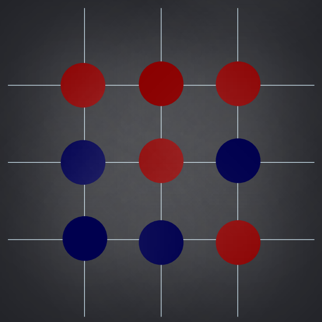


Spin models

spin up $\sigma = +1$

spin down $\sigma = -1$

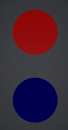
Lattice
size = L x L



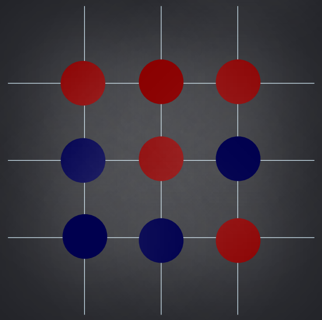
Spin models

spin up $\sigma = +1$

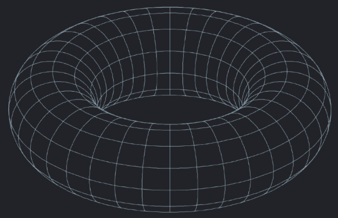
spin down $\sigma = -1$



Lattice
size = L x L



Periodic
boundary
conditions

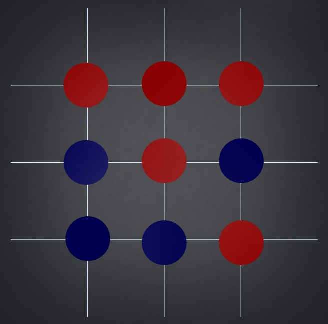


Spin models

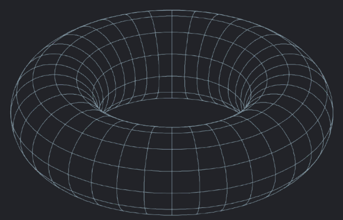
spin up $\sigma = +1$

spin down $\sigma = -1$

Lattice
size = L x L



Periodic
boundary
conditions



Ising model[1]

- Square lattice □
- Interacts 4 neighbors

$$H_{is} = -\frac{J}{2}\sum_{\langle i,j\rangle} \sigma_i \cdot \sigma_j$$

Baxter-Wu model[2]

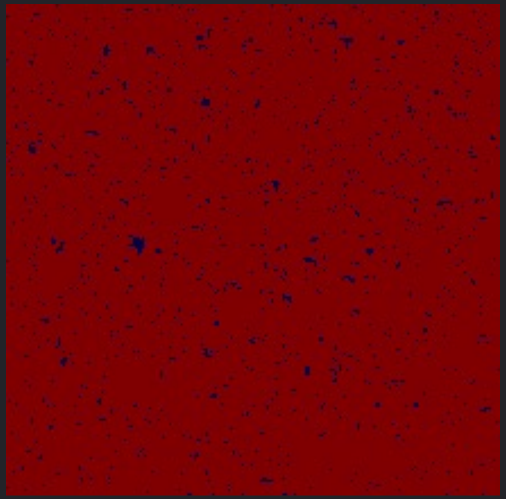
- Triangular lattice △
- Interacts 6 neighbors

$$H_{bw} = -J \cdot \sum_{\langle faces\rangle} \sigma_i \cdot \sigma_j \cdot \sigma_k$$

1 Lars Onsager. "Crystal statistics. I. A two-dimensional model with an order-disorder transition". In: Physical Review 65.3-4 (1944), p. 117.
2 Rodney J Baxter and FY Wu. "Ising model on a triangular lattice with three-spin interactions. I. The eigenvalue equation". In: Australian Journal of Physics 27.3 (1974), pp. 357–368.

Phase transition

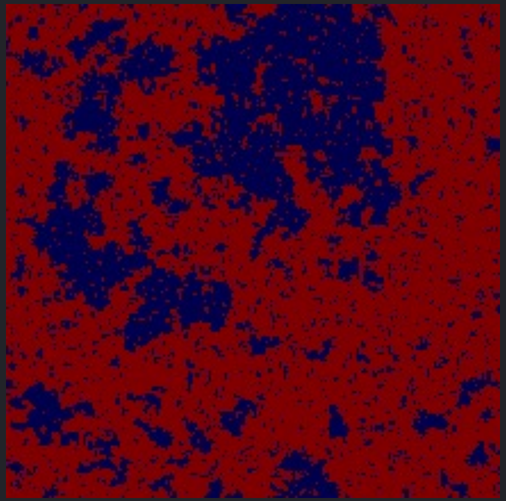
Ferromagnetic phase



$T = 1.869$

Low-temperature

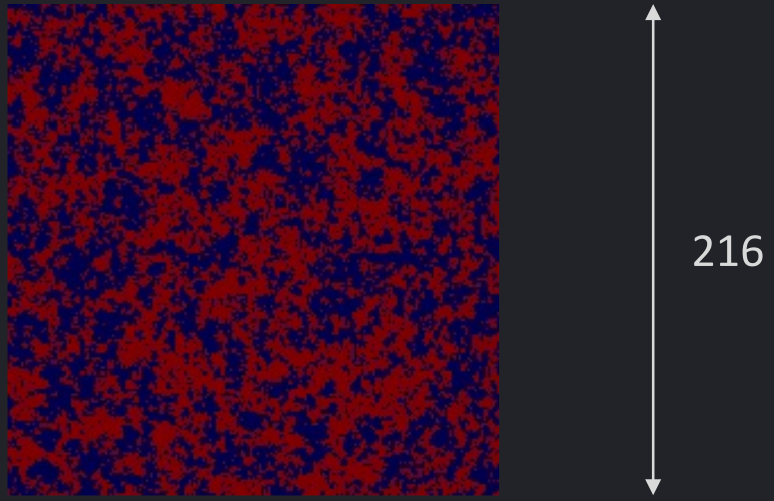
Transition point



$T_c = 2.269$

Critical temperature

Paramagnetic phase



$T = 2.719$

High-temperature

Generate uncorrelated data

Monte Carlo is an extremely bad method; it should be used only when all alternative methods are worse.[1] *© A. Sokal*

> $error \sim 1/\sqrt{n\_iter}$

[1] Alan Sokal. "Monte Carlo methods in statistical mechanics: foundations and new algorithms". In: Functional integration. Springer, 1997, pp. 131–192.

Generate uncorrelated data

Monte Carlo is an extremely bad method; it should be used only when all alternative methods are worse.[1] *© A. Sokal*

$$error \sim 1/\sqrt{n\_iter}$$

Metropolis Monte Carlo (single spin flip)

initialize spins

repeat N_{flip} times:

pick random spin

if $\Delta E < 0$ → update

else if $\exp(-\Delta E/T) \geq rnd()$ → update

Let $L = 243,\ \ N_T = 126$

$N_{img} = 189\,000$ (1500 per T)

$$N_{flip} = 20 \cdot t_{corr} \cdot N_T + 2\, t_{corr} \cdot N_{img} \approx 3 \cdot 10^{15}$$

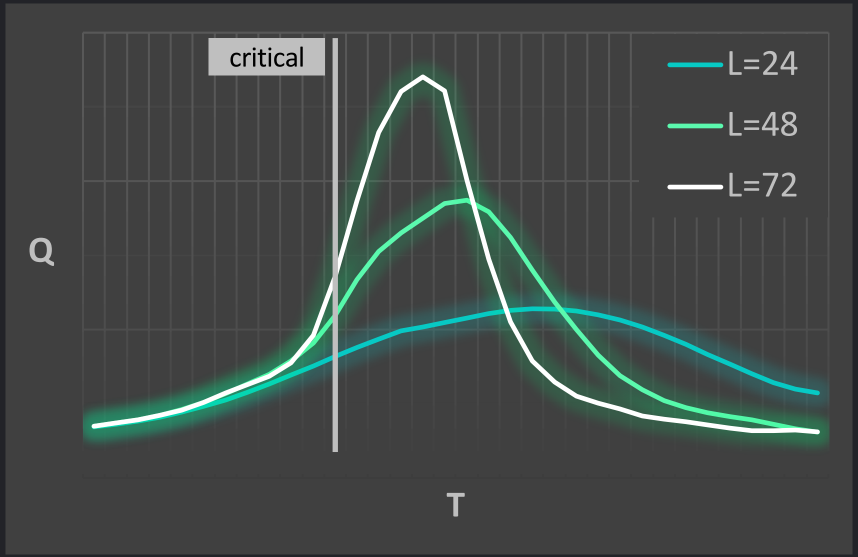
$$t_{corr} = L^2 \cdot L^{2.15}$$

[1] Alan Sokal. “Monte Carlo methods in statistical mechanics: foundations and new algorithms”. In: Functional integration. Springer, 1997, pp. 131–192.

Conventional method

Finite-size scaling (FSS)

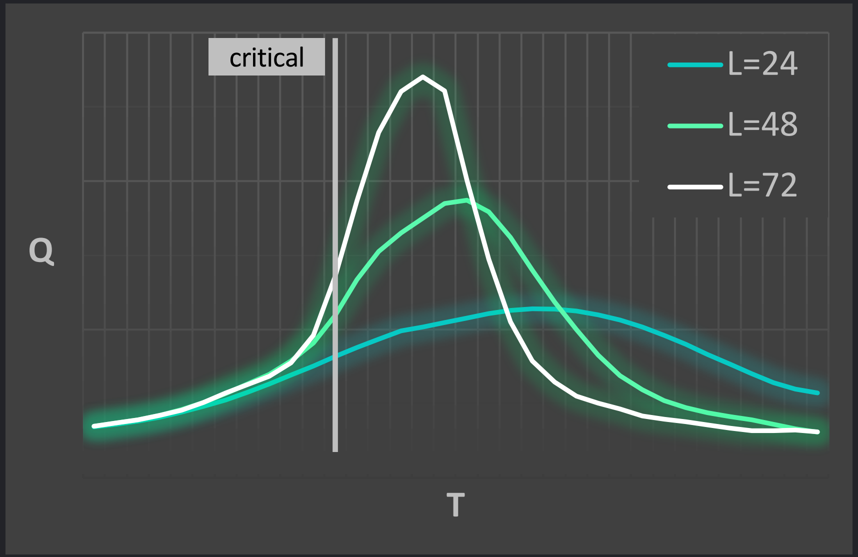
Typical thermodynamic quantity **Q(T)** scales when system size **L** increases



Conventional method

Finite-size scaling (FSS)

Typical thermodynamic quantity **Q(T)** scales when system size **L** increases



Extract critical exponent

Finite-size scaling (FSS) of thermodynamic quantities

Model	**Universality class**	$\boldsymbol{\alpha}$ from C	$\boldsymbol{\beta}$ from M	$\boldsymbol{\gamma}$ from $\boldsymbol{\chi}$	$\boldsymbol{\nu}$ any
Ising	Ising	0	1/8	7/4	1
Baxter-Wu	4-st. Potts	2/3	1/12	7/6	2/3

Recent advances

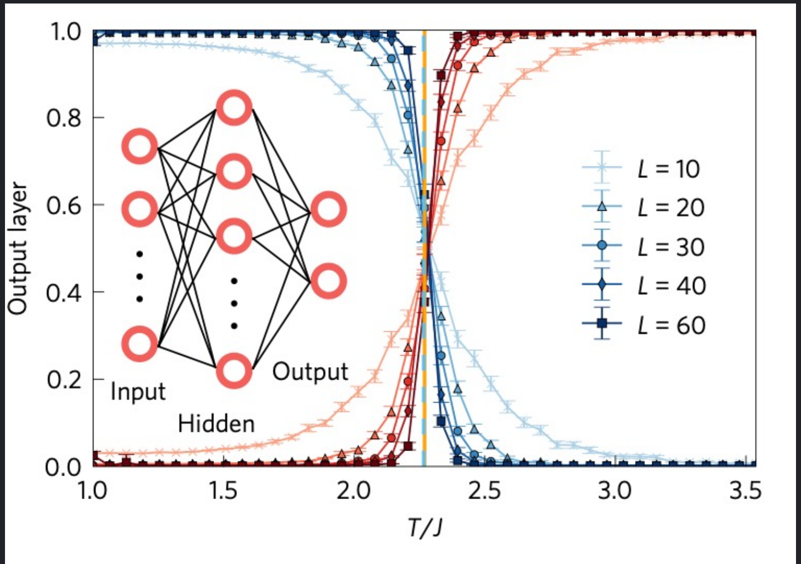
In “Machine learning phases of matter[1]” applied neural network (NN) to predict phase of spin configuration (image).

[1] Juan Carrasquilla and Roger G Melko. “Machine learning phases of matter”. In: Nature Physics 13.5 (2017), pp. 431–434.

Recent advances

In “Machine learning phases of matter [1]” applied neural network (NN) to predict phase of spin configuration (image).

trained & predicted $T_c = 2.266(2)$



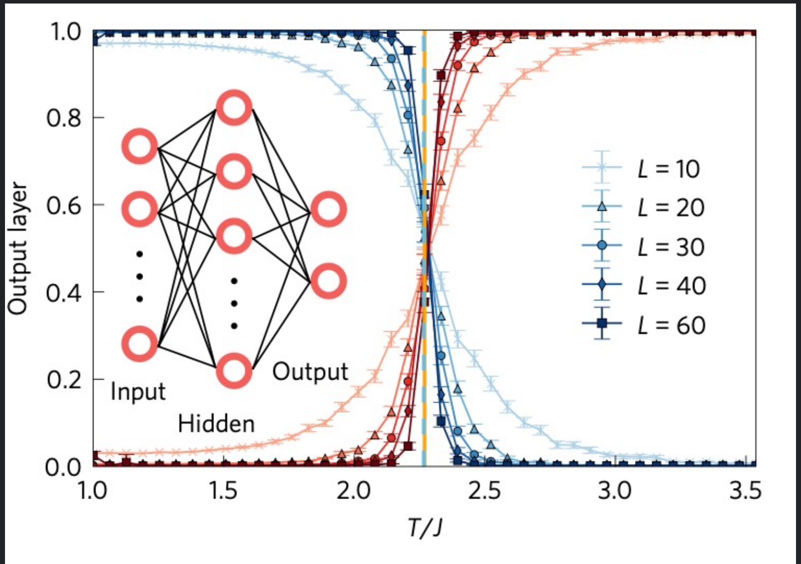
2D Ising [square]

[1] Juan Carrasquilla and Roger G Melko. “Machine learning phases of matter”. In: Nature Physics 13.5 (2017), pp. 431–434.

Recent advances

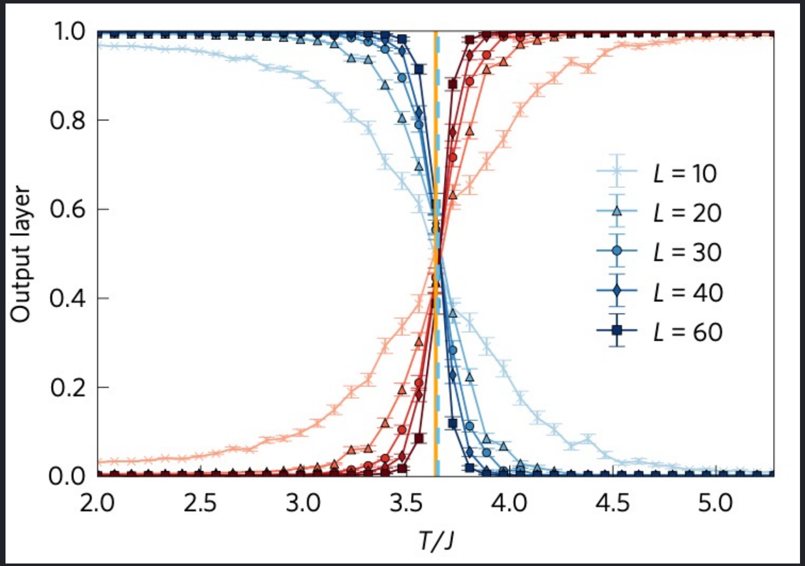
In "Machine learning phases of matter[1]" applied neural network (NN) to predict phase of spin configuration (image).

trained & predicted T_c = 2.266(2)



2D Ising [square]

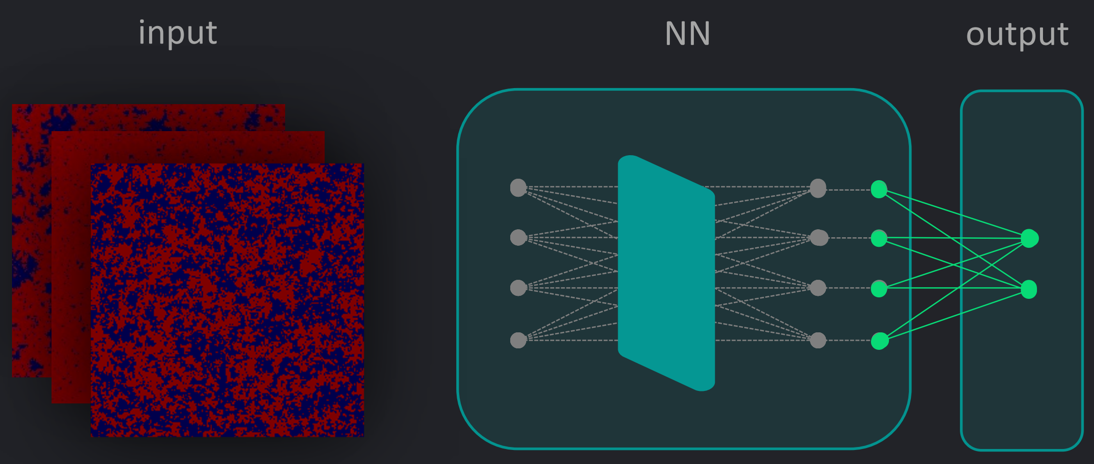
predicted T_c = 3.65(1)



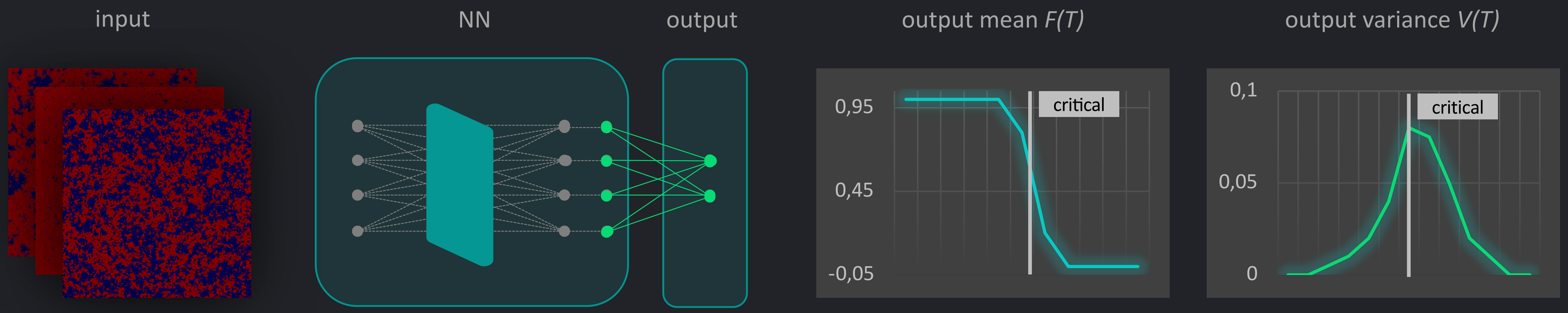
2D Ising [triangular]

[1] Juan Carrasquilla and Roger G Melko. "Machine learning phases of matter". In: Nature Physics 13.5 (2017), pp. 431–434.

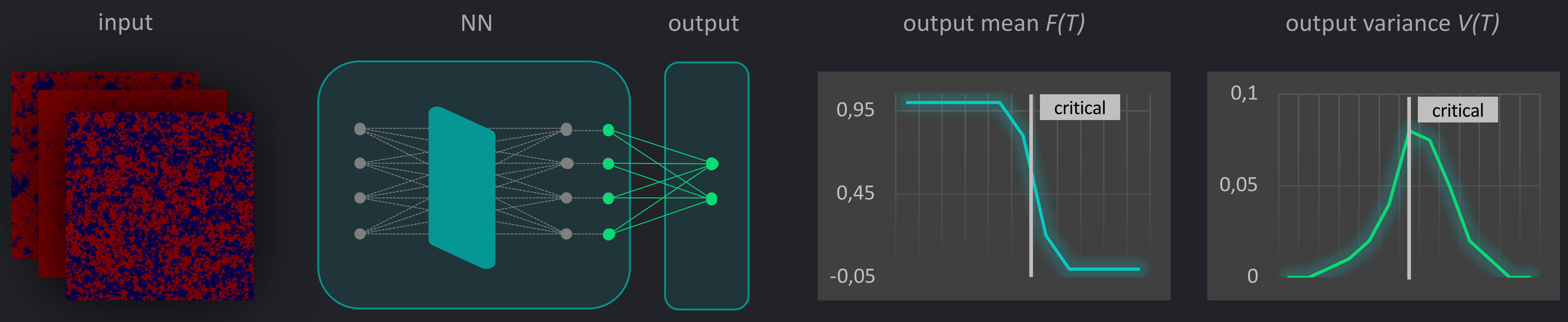
NN method



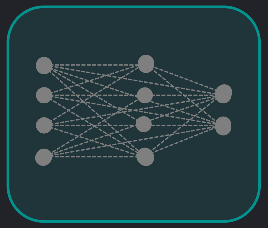
NN method


input
NN
output
output mean F(T)
0,95
0,45
-0,05
critical
output variance V(T)
0,1
0,05
0
critical

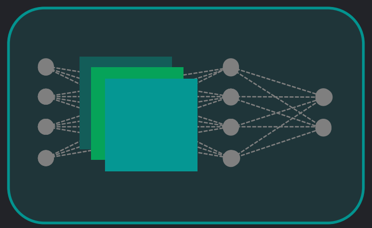
NN method



Architectures



FCNN

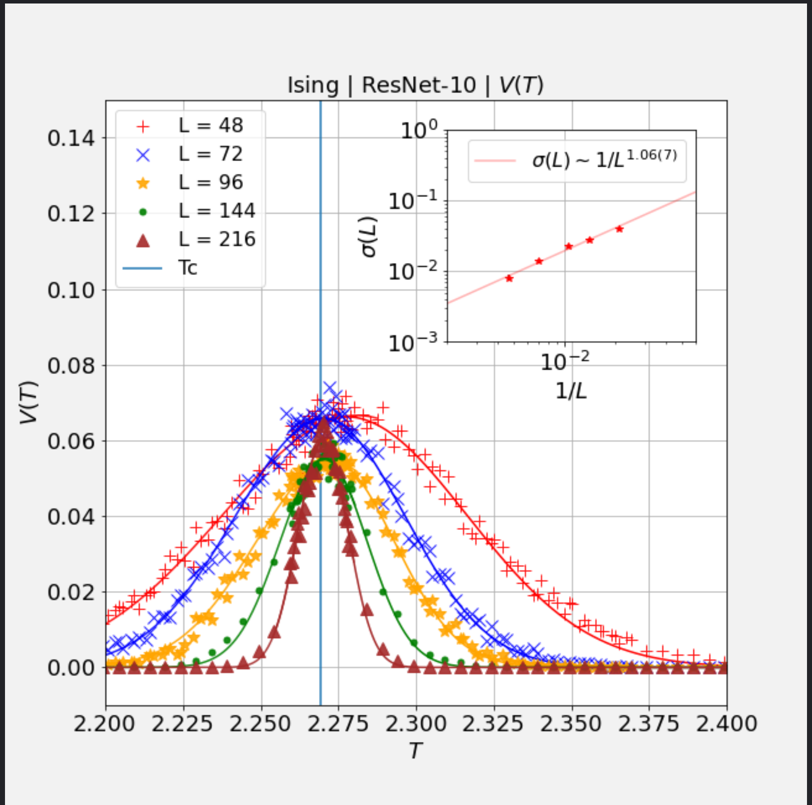


ConvNN

ResNet[1] family (10, 18, 34, 50 layers)

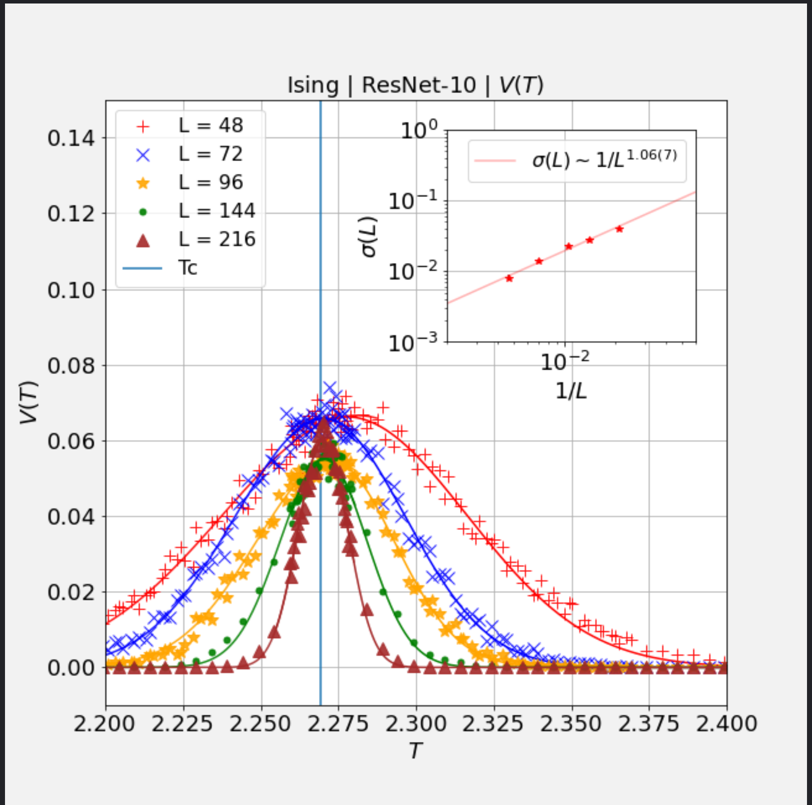
[1] Kaiming He et al. “Deep residual learning for image recognition”. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 770–778.

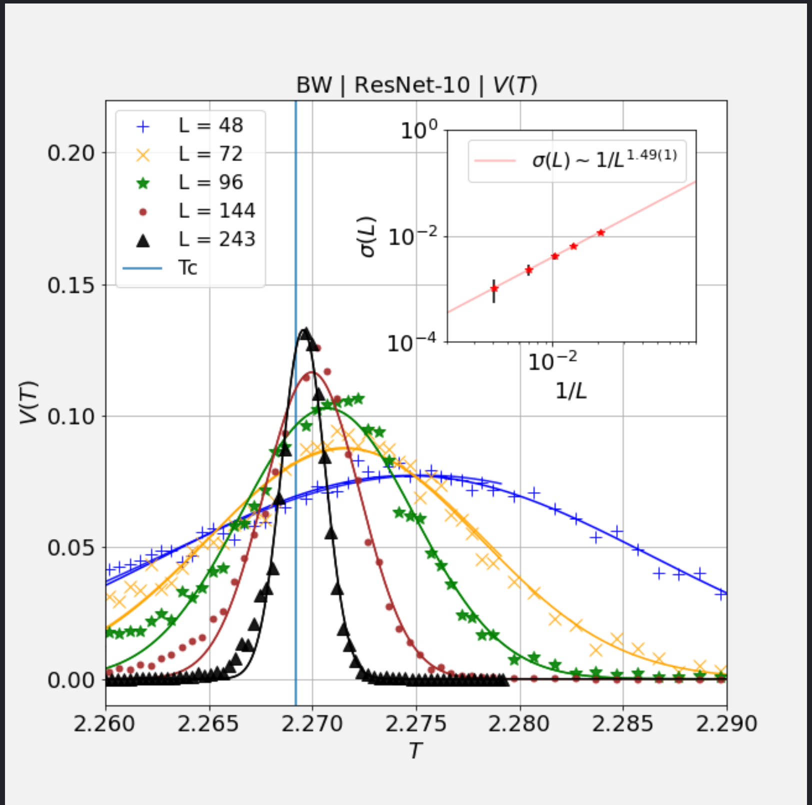
Exponents estimation



Model	$1/\nu$ theoretical	$1/\nu$ conventional	$1/\nu$ NN method
Ising	1	1.02(5)	1.06(7)

Exponents estimation





Model	$1/\nu$ theoretical	$1/\nu$ conventional	$1/\nu$ NN method
Ising	1	1.02(5)	1.06(7)

Model	$1/\nu$ theoretical	$1/\nu$ conventional	$1/\nu$ NN method
Baxter-Wu	1.5	1.52(3)	1.49(1)

Depth dependence

Models

Model	#params, 10^6
ResNet-10	4.9
ResNet-18	11.2
ResNet-34	21.3
ResNet-50	23.5

Depth dependence

Models

Model	#params, 10^6
ResNet-10	4.9
ResNet-18	11.2
ResNet-34	21.3
ResNet-50	23.5

Ising

$\mathbf{1/\nu}$ theoretical	$\mathbf{1/\nu}$ NN method
1	1.06(7)
	1.03(7)
	1.07(8)
	1.09(11)

Baxter-Wu

$\mathbf{1/\nu}$ theoretical	$\mathbf{1/\nu}$ NN method
1.5	1.49(1)
	1.50(3)
	1.50(2)
	1.52(1)

Conclusion

- Estimate critical exponents ν for both models with the same accuracy using conventional (FSS) & NN methods.
- No evidence that the quality ν extraction depends on the number of convolutional layers (different ResNet-s).
- Fluctuation of the NN output as a function of temperature has a characteristic Gaussian shape.
- NN learns the location of the phase transition, critical exponent ν of the universality class of the model.

Outlook

- Transfer learning: whether and to what accuracy an NN trained on one model, predicts critical properties of a different model in the same universality class?
- Whether NN learns only the correlation length exponent ν, or if other critical exponents can be extracted from the NN outputs?

Acknowledgements

This work is supported by the grant 22-11-00259 of the Russian Science Foundation. Simulations were done using the computational resources of HPC[1] facilities at HSE University.

[1] PS Kostenetskiy, RA Chulkevich, and VI Kozyrev. “HPC resources of the higher school of economics”. In: Journal of Physics: Conference Series. Vol. 1740. 1. IOP Publishing. 2021, p. 012050.

Appendix

Samples generating

Intel Xeon Gold 6152

Time required to generate data for Ising model:

Size	Total single CPU time, hour	Real time, hour
48	65	0.5
72	354	2.9
96	1071	8.5
144	6098	48
216	20492	162

NN training

NVIDIA Tesla V100-SXM2 32 GB

Training time for one epoch, Ising model, L=48:

NN type	#params, 10^6	Time, s/epoch
ConvNN	0.59	108
FCNN	0.23	66
ResNet-10	4.9	534
ResNet-18	11.2	1200
ResNet-34	21.3	2369
ResNet-50	23.5	2590