Machine Learning models for identifying phase transitions

Pratyush Jha (20410022) | Engineering Physics

1.Introduction:

In condensed matter physics, the size of the state space, which grows exponentially as the number of particles, governs the complexity of the problem. This situation, in machine learning, is reminiscent of the 'curse of dimensionality' issue [1] which leads to incorrect predictions by the model due to overfitting. Despite this issue, multiple machine learning architectures have been developed that perform remarkably great in their abilities to recognize and correctly classify complex data. In this project, the ability of neural networks to correctly identify phases and phase transitions have been explored. We observe that machine learning architectures such as a standard feed-forward neural network and a convolutional neural network [2][3] can identify such phase transitions without any knowledge of the Hamiltonian involved or the general locality of interactions. The goal of my project was to study the phase transition in Ising model and apply machine learning models (feed-forward neural network and convolutional neural network) for identifying phase transitions by reproducing some of the results published in Ref. [4]. The code for the same can be found here.

2.Using a fully connected neural network for the Ferromagnetic Ising Model:

A feed-forward neural network with a hidden layer consisting of 100 neurons classifies the given input state as a ferromagnetic or paramagnetic state with a very high accuracy as expected from the Ising model.

2.1.The Ferromagnetic Ising model:

The Ising model in statistical mechanics describes the behaviour of a collection of interacting spins on a lattice. Consider a square lattice of dimensions $N \times N$. A magnetic spin(σ_i), called the Ising spin, is present at each lattice point. This system has two degrees of freedom such that each spin value can be considered in two possible states: up (+1) or down (-1). Hence, it can be observed that for N lattice sites, the state space is of the size 2^N. In this model, energy includes contributions from neighbouring spin interactions and the effect of externally applied magnetic field on every spin. This can be summarized as:

$$H = -\sum_{<i,j>} J_{ij}\sigma_i^z\sigma_j^z - h\sum_i \sigma_i$$

Where, $< i, j >$ refers to the nearest-neighbour pair and J_{ij} is the exchange energy term. We will consider $J_{ij} \equiv J$, i.e., the exchange energy term as a constant in all cases.

The first term of the Hamiltonian depicts interaction between spins present at the lattice points whereas the second term corresponds to the energy due to an externally applied magnetic field.

We can drop the second term of the Hamiltonian by assuming the absence of any applied magnetic field.

It can be observed that the Ising model Hamiltonian favours all the spins to align in the same direction to get to the lowest energy state. This is true at absolute zero. However, for a two-dimensional lattice following the Ising model, at temperatures above absolute zero, the average magnetisation decreases until it reaches zero at Curie Temperature [5]. This is because at temperatures greater than the Curie Temperature, thermal motion disrupts the alignment. Hence, we can define two states: a ferromagnetic state below the Curie Temperature and a paramagnetic state above it. Thus, the lattice is said to undergo a second order phase transition from an ordered phase to a disordered phase at the Curie Temperature.

The Curie Temperature under the thermodynamic limit for a two-dimensional Ising model with $H = -J \sum_{<i,j>} \sigma_i^z \sigma_j^z$ is found to be:

$$\frac{kT_c}{J} = {}^{2}\!/_{\ln(1+\sqrt{2})}$$

Where, T_c is the Curie Temperature and k is the Boltzmann constant.

In our discussions, we will set $J = 1$ and will consider only the first term of the Hamiltonian due to the absence of any externally applied magnetic field.

2.2.Dataset generation:

The Standard Monte Carlo technique is used to get the sample configurations for an array of temperature values T, weighted by the Boltzmann distribution. These sample configurations represent our lattice in the form of a matrix with each matrix element representing the spin at a lattice point.

A total of 26000 sample configurations were generated for lattice size values of N = 10, 20 and 30. For each value of N, 85% of the generated data is for training the model while the rest is used as test dataset.

2.3.Defining the neural network architecture:

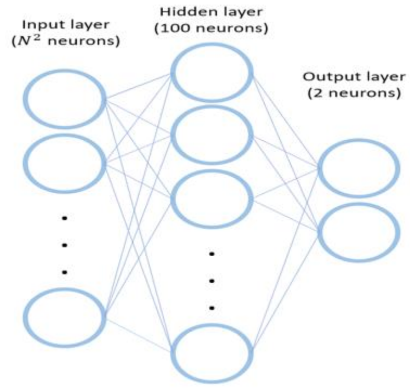


Fig. 1: Illustration of the neural network model

The neural network is defined with an input layer (which has a flattened vector comprising of total N^2 values for the $N \times N$ lattice space), a 100-unit hidden layer and a 2-value output layer that gives the probability of the configuration to be in ferromagnetic state and that of it to being in the paramagnetic state, thereby, classifying between the two.

2.4.Training the neural network:

The sample lattice configurations prepared as per the algorithm discussed in section 2.2 are fed to the neural network as the dataset. Since the Critical temperature for phase transition is known to be T_c, we can classify our dataset into 2 categories: the *lower-temperature ferromagnetic phase* and the *higher-temperature paramagnetic phase*. Our fully connected feed-forward neural network trains on the provided training dataset to learn this classification task.

2.5 Results:

2.5.1. On training the neural network with training data corresponding to N = 10, 20 and 30, we notice that the model is able to better learn the phase transition when the input lattice size is larger (Refer to Fig. 2).

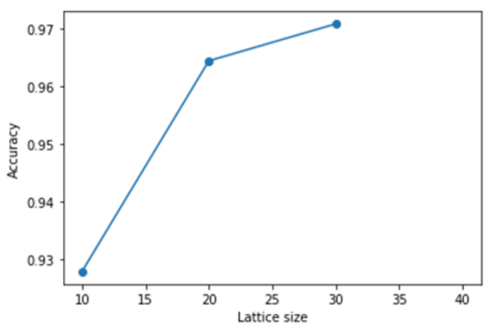


Fig. 2: Accuracy of trained model on the test dataset with variation in lattice size (N= 10, 20 and 30).

2.5.2. The model is fairly but not perfectly accurate with respect to its predictions when the temperature of the lattice is around the Curie Temperature. This can be observed from the graphs plotted in fig. 3 in which the output layer values are not very close to 0 or 1, showing that the model is discriminating the transition with a lesser accuracy. This issue is more relevant when the input configuration is of lower dimensions, as can be observed by comparing figures 3(a), 3(b) and 3(c).

2.5.3. The Curie Temperature value learnt by the model is very close to the value of $^2/_{\ln{(1+\sqrt{2})}}$ in the thermodynamic limit for a square lattice. This is depicted in Fig. 3. The green line in figures 3(a), 3(b) and 3(c) depict the expected value, whereas the point of intersection of the output values of the two layers (shown in blue and orange) is the value of Curie Temperature learnt by the model. This is also shown that the Curie Temperature value learnt by the model is more precise in case of N=30 and N=20 than N=10 where N is the input lattice size value for an $N \times N$ lattice as the intersection point of the two curves is observed to be closer to the green line.

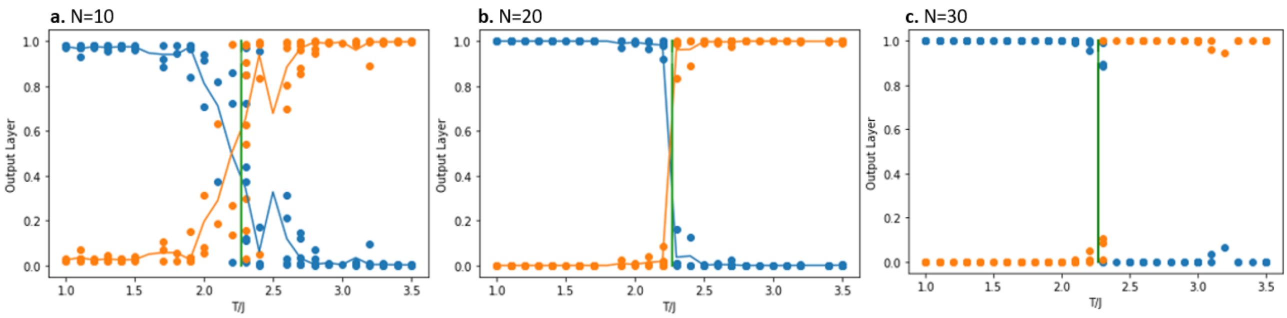


Fig. 3: The blue line depicts probability of input being ferromagnetic whereas the red line depicts probability of input being paramagnetic with temperature on the X-axis for models trained with (a.) N=10, (b.) N=20 and (c.)N=30.

The green line depicts the value corresponding to $\frac{kT_C}{J} = {}^2/_{\ln(1+\sqrt{2})}$ whereas the intersection of blue and orange curve is the model's prediction of the Curie Temperature.

These predictions are on the test dataset, thus depicting that the model is able to generalize the results well.

It is worth noting that we also get results with similar accuracy values while considering anti-ferromagnetic couplings with the Hamiltonian as $H = J\sum_{<i,j>}\sigma_i^z\sigma_j^z$.

3.Using a convolutional neural network (CNN) for identification of phase transitions:

Convolutional Neural Networks emerged as a major breakthrough in processing information in the form of a matrix. In case of computer vision tasks, this data is generally a matrix of pixel values of the image. Since our data is also in the form of a matrix, we utilized this by using a CNN for the classification task and compared the results.

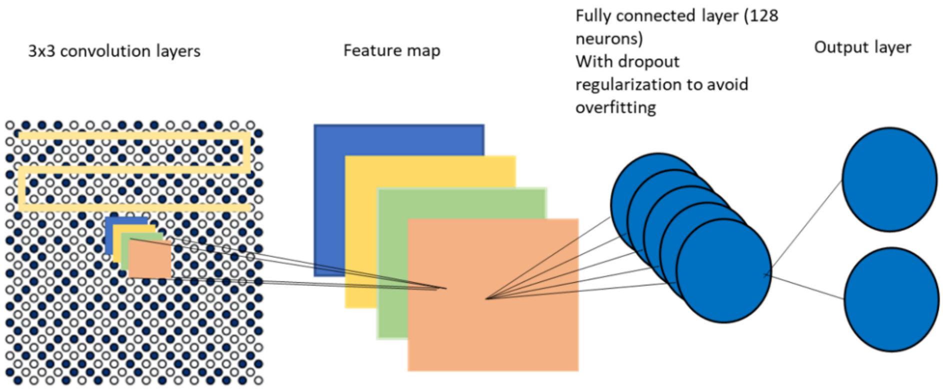


Fig. 4: Illustration of the CNN model

The input lattice configuration is convolved with a 3x3 kernel using the ReLu (Rectified Linear Unit) activation function. This is followed by a two fully connected layers getting the input as an $N^2 \times 1$ dimensional flattened matrix whose output is the softmax probability of the sample configuration being ferromagnetic and paramagnetic, as in the case of feed-forward network. The Convolutional Neural Network was able to perform at a very high accuracy of more than 99% in the case of Ising model, as compared to the ~97% accuracy in case of Feed-forward neural networks. This is because of the ability of CNN to take benefit of the translational invariance and preserve and learn from the spatial information of the lattice.

3.1 CNN for Ising Lattice Gauge Theory:

The importance of CNN increases all the more when we consider the classification task in case of Ising lattice gauge theory as discussed below:

3.1.1 Ising Lattice Gauge Theory:

In Ising lattice gauge theory, spinors are present on the links joining the lattice sites. In this case, the Hamiltonian thus obtained is given as:

$$H = -J\Sigma_p \Pi_{i\in p}\sigma_i^z$$

Where p refers to the plaquettes.

3.1.2 Using CNN for phase transitions in Ising Lattice Gauge Theory:

Using a neural network for classification purpose for $T = 0$ and $T \to \infty$ states in Ising Lattice Gauge Theory is not useful due to the increased importance of the spatial structure.

Since the neural network requires a flattened $N^2 \times 1$ vector as an input for the corresponding $N \times N$ lattice, it loses the spatial information of the lattice required for correct classification. Hence, a better choice in such a case is a Convolutional Neural Network (CNN) which preserves this information. A very high accuracy of upto 100% can be achieved using a CNN as described in Ref. [4].

4.Conclusion:

The project shows that machine learning models such as feed-forward neural networks and CNNs have the ability to very accurately encode phases of matter along with the ability to correctly discriminate phase transitions in correlated many-body systems without any knowledge of the Hamiltonian. Along with this, Ref. [4] discusses that they have the ability to encode basic information about unconventional phases as in the case of Ising Lattice Gauge Theory and the square-ice model.

Thus, machine learning techniques, such as neural networks and convolutional neural networks (CNN), are helpful in the identification of phase transitions in condensed matter physics due to their ability to identify complex patterns and relationships within data. In condensed matter physics, phase transitions often involve changes in the statistical properties of materials, which can be difficult to identify using traditional methods. Machine learning algorithms can be trained on large datasets of condensed matter systems to recognize the patterns that are characteristic of different phases.

Once trained, these algorithms can accurately classify the phases of new materials based on their statistical properties. This can help researchers identify new phases of matter, predict phase transitions under different conditions, and understand the underlying physics behind phase transitions. Furthermore, machine learning algorithms can process large amounts of data quickly and efficiently, making it possible to analyse complex systems that would otherwise be too time-consuming for traditional methods.

Overall, machine learning has the potential to significantly enhance our understanding of condensed matter physics and contribute to the development of new materials with unique properties.

References:

[1] Bellman, R. Dynamic Programming 1st edn (Princeton Univ. Press, 1957).

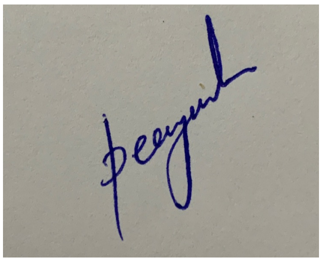
[2] Neupert, T. et al. (2022) Introduction to machine learning for the Sciences, arXiv.org. Available at: https://arxiv.org/abs/2102.04883.

[3] Goodfellow, I., Bengio, Y. & Courville, A. Deep Learning (MIT Press, 2016); http://www.deeplearningbook.org

[4] Carrasquilla, J., Melko, R. Machine learning phases of matter. *Nature Phys* **13**, 431–434 (2017). https://doi.org/10.1038/nphys4035

[5] Study of the ising model - web.njit.edu (no date). Available at: https://web.njit.edu/~kenahn/SEED/2009/final_project_Shenelle.pdf

(Pratyush Jha)
Student’s signature



(Peayush Kumar Choubey)
Supervisor’s signature