Identification of the Critical Temperature for Spin Models Using Machine Learning

Sam Foreman

University of Iowa

samuel-foreman@uiowa.edu

February 13, 2017

Overview

Motivation

- Can machine learning be used to identify phase transitions in spin systems?
- How can we generalize this model to systems with more complicated structure?
  - e.g. $O(2)$ model, topological phases, detection of order parameters etc.
- Useful as complementary method to verify Monte Carlo results.
- Problems with current approaches:
  - Monte-Carlo methods are often computationally exhaustive.
  - The sign problem when dealing with high-dimensional fermion systems.

Motivation (cont'd)

- We show that convolutional neural networks (CNN) are capable of accurately predicting the critical temperature of the phase transition in both the 2D Ising and XY spin models.
- Using spin configuration data generated from Monte-Carlo sampling, we implement state-of-the-art machine learning algorithms to classify the systems phase.
- We begin by performing a thorough analysis on the solvable 2D Ising model, and attempt to generalize these results to the 2D XY model.

Setup

- Begin with simple 2D spin systems:
  - 2D Ising Spin Model, with

$$H_{ising}(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

where $\sigma_i \in \{-1, +1\}$

  - 2D XY Spin Model, with

$$H_{XY}(\theta) = -J \sum_{i \neq j} cos\left(\theta_i - \theta_j\right)$$

where $\theta_i \in (-\pi, \pi)$

Setup

- Use Metropolis algorithm to generate example configuration data.
- Perform both supervised and unsupervised learning methods to locate a phase transition.
- Supervised learning is useful when we are able to provide labeled data (e.g. spins generated from Monte-Carlo techniques) consisting of a configuration, along with a phase label i.e. $T \lesseqgtr T_c$.
- Unsupervised learning is useful when we have unlabeled data, or when we wish to validate the results obtained from supervised learning.

Setup

- **Supervised Learning:**
  - For the Ising model, we expect a phase transition at

$$T_c = \frac{2J}{k_B \ln(1+\sqrt{2})} \approx 2.269 J/k_b$$

  - For the XY model, we also expect a phase transition at a temperature near $T_c \approx 0.893$.
- **Unsupervised Learning:**
  - For the Ising model, we also perform unsupervised learning, where we perform a cluster analysis on dimensionally-reduced configuration data that is able to identify unique phases without being given prior information about the configurations temperature.

Neural Network

- Perform supervised learning with input data consisting of N spin configurations on an $L \times L$ lattice, at a specific temperature T.

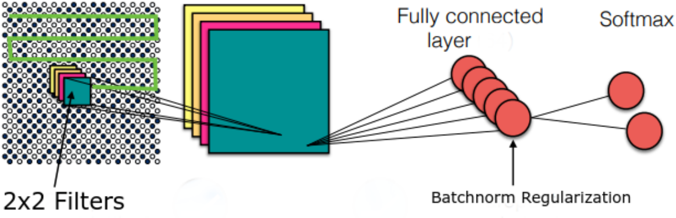
$$\left\{\left(\left\{\sigma_{ij}^{n}\right\}, T_n\right) | T_n = T_0 + n\epsilon\right\}_{n=0,\dots,N} \tag{1}$$

- Create binary training labels to identify the phase:

$$y(T_n) = \begin{cases} 0, & \text{if } T_n < T_c \\ 1, & \text{if } T_n > T_c \end{cases}$$

Neural Network (cont'd)

- Construct convolutional neural network, with architecture[1]



- Convolutional neural networks have the advantage of using a coarse-grained (hierarchical) approach to help locate features for a given configuration.
- This approach is reminiscent of the 'block variables' used in renormalization group techniques, which help to separate internal and external degrees of freedom.
- During a given forward pass, a filter is convolved across the input data, producing an activation map which gives the responses at every spatial position.

[1] arXiv:1605.01735 [cond-mat.str-el]

Neural Network (cont'd)

- We use a softmax classifier as the output layer, which computes a probability distribution over the possible outcomes.

$$P(y = j|\mathbf{x}) = \frac{e^{\mathbf{x}^T \mathbf{w}_j}}{\sum_{k=1}^{K} e^{\mathbf{x}^T \mathbf{w}_k}} \tag{2}$$

  where $\mathbf{x}^T \mathbf{w}_j$ is the j^{th} component of the previous layers output.

- Since we are using a binary classification system (i.e. $T \lessgtr T_c$), the network will output an array of scores,

$$S = [s_{below}, s_{above}]$$

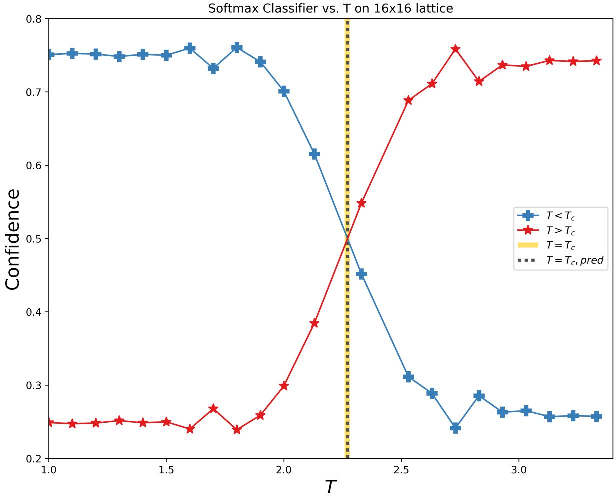
  where

$$s_{below} + s_{above} = 1$$

- We can interpret this output as a confidence measure of its predicted classification.

Supervised Learning on Ising Model

Results after training for 10 epochs.



Results (cont'd)

- From this result, we see that our network correctly classifies configurations at temperatures well above and below the critical temperature.
- At temperatures $T \approx T_c$, we see that this classification scheme breaks down.
- This signifies a hidden structure (feature) in our data, which we know to be representative of a phase transition.

Unsupervised Learning on Ising Model

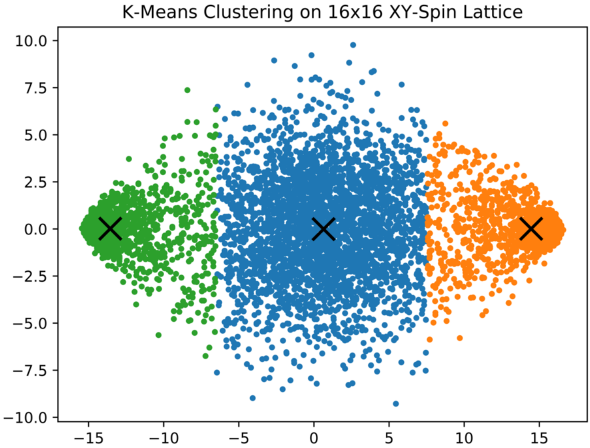
- Principal Component Analysis (PCA):
  - Used to project configuration data onto the dimensions with the largest variances.
  - This is accomplished by performing SVD decomposition on configuration data.
  - Applied to Ising spins, PCA finds the most significant variations of the data changing with temperature.
  - We expect that this principal component should be representative of the uniform magnetization,

$$m = \frac{1}{N}\sum_i \sigma_i \tag{3}$$

- k-Means Clustering:
  - Having projected each configuration onto this low-dimensional space, we then wish to partition these configurations into k unique clusters.

k-means Clustering Results

- After projecting the covariance matrix onto the two largest eigenvalues, we can clearly identify three distinct regions:



Supervised Learning on XY Model

- Results after training for 10 epochs on 32×32 XY spin lattice.

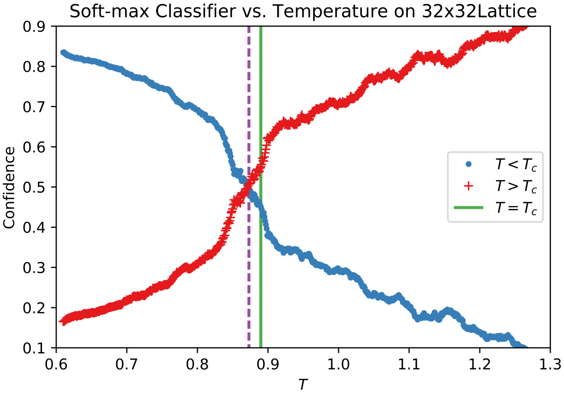


Figure: Note that the green line is an approximation to the predicted critical temperature, $T_c \approx 0.893$, and the purple line represents our networks predicted critical temperature, $T_{c,pred} = 0.873$

- Supervised Learning
  - We have shown that standard neural networks are capable of not only recognizing hidden features, but are also able to provide a meaningful insight into physical parameters that govern the phase transition in complex spin systems.
  - By employing both supervised and unsupervised learning algorithms on raw configuration data, we confirm that our model correctly distinguishes between phases, even without being given any *a priori* information about the underlying physical structure of the system.

References

1. arXiv:1605.01735 [cond-mat.str-el]
2. arXiv:1609.09087 [cond-mat.dis-nn]
3. arXiv:1606.00318 [cond-mat.stat-mech]
4. arXiv:1608.07848 [cond-mat.str-el]