DISCOVERY PHSE TRANSITIONS USING UNSUPERVISED MACHINE LEARNING PCA

Jairo Jose Orozco Sandoval

Advisor:

Ph.D. Junqiang Lu

University of Puerto Rico – Mayagüez campus
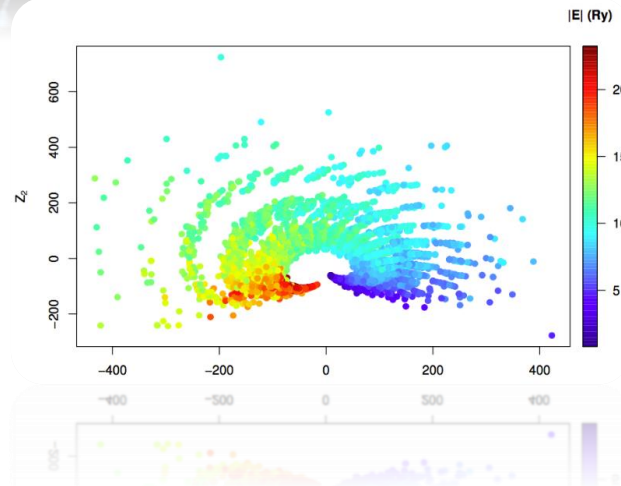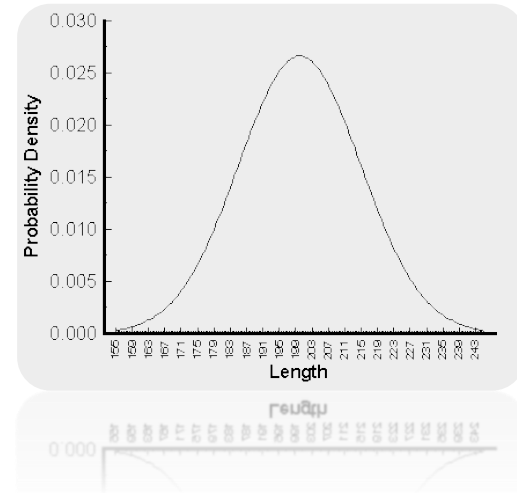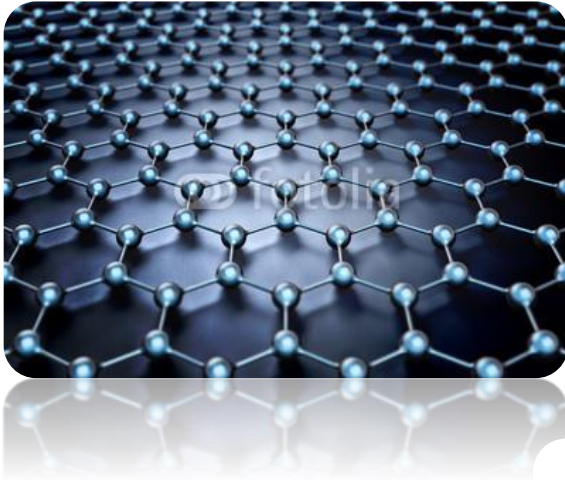April 25 2019

Content

- Introduction

- Objectives

- The Ising model

- Monte Carlo method
  - Metropolis Algorithm

- Learning Algorithms
  - Principal Component Analysis (PCA)

- Results

- Conclusions

Introduction

Machine learning

Introduction

Machine learning

Supervised Learning

Unsupervised Learning

Phase transition

ferromagnetism

paramagnetism
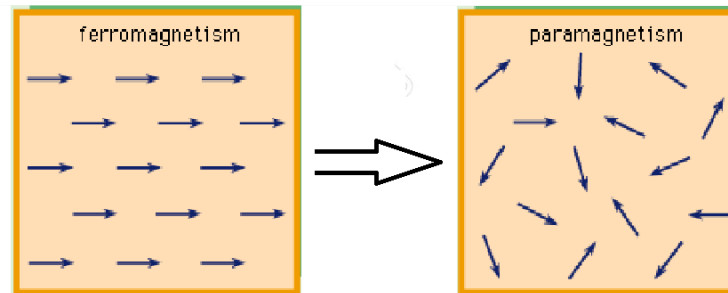
Objectives

- Implement machine learning techniques to identify phases and phase transitions in 2D Square Ising model

- Use machine learning to find order parameters, distinguish phase transition and locate critical temperature.

General idea

- We apply unsupervised machine learning technique, the Principal Component Analysis (PCA) to study square ferromagnetic Ising model with lattice sizes L= 10,20,30,40 ,50

The Ising Model

- Was proposed by Lenz in 1920

- 1925 was discussed by Ernst Ising .

- The exact solution of the two-dimensional Ising model was made by Onsager in 1944
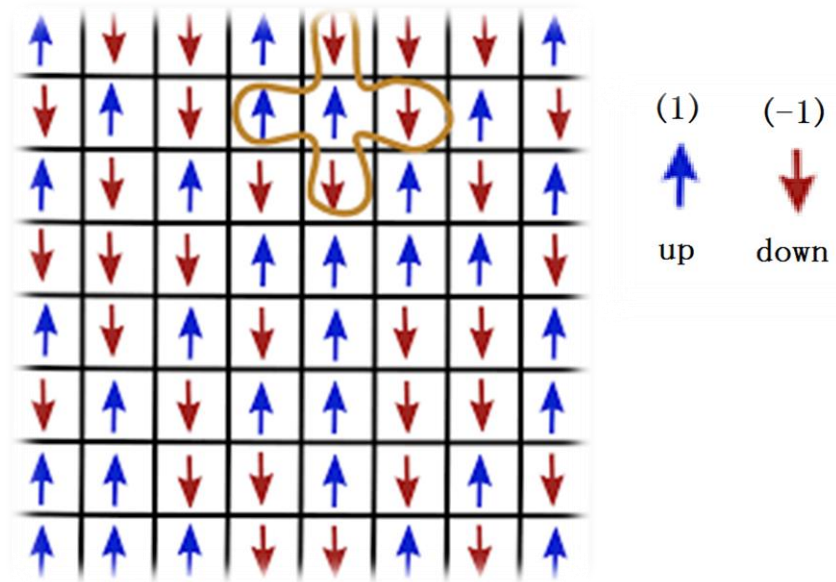
$$H = -J \sum_{<i,j>} S_i S_j - h \sum_i S_i$$

$$T_c = 2J/K_B \ln(\sqrt{2} + 1) = 2.2692(J/K_B)$$

$J > 0 \quad ferromagnetic$

$J < 0 \quad antiferromagnetic$

The Ising model:
Phase transition and finite size scaling

- In the Ising model the 1st derivative of the free energy gives the magnetization M and the 2nd derivative gives the magnetic susceptibility χ_m

$$M = \frac{1}{V}\left(\frac{dF}{dH}\right)_{h=0} \quad \chi_m = \frac{1}{V}\left(\frac{d^2 F}{dH^2}\right)_{h=0}$$

- Magnetization, $M \sim |T - T_C|^\beta$

- Magnetic susceptibility, $\chi \sim |T - T_C|^{-\gamma}$

- Heat capacity, $C \sim |T - T_C|^{-\alpha}$

- Correlation length, $\xi \sim |T - T_C|^{-\nu}$

The critical exponents for the 2D Ising model are known exactly:

$$\beta = 0.125 \qquad \alpha = 0 \qquad \gamma = 1.75 \qquad \nu = 1$$

The Ising model:
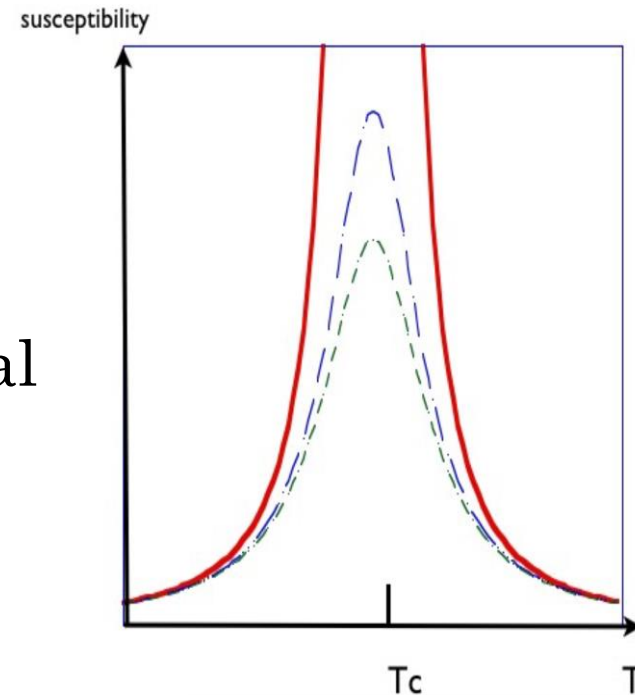## Phase transition and finite size scaling

- In an infinite system ($L \sim \infty$)

- In a finite size the correlation length is $\xi \sim L$.

- The system has a pseudocritical point when

$$\left| T_{C(\infty)} - T_{C(L)} \right|^{-\nu} \sim L$$

- Then the susceptibility,

$$\chi \sim L^{\gamma/\nu}$$

susceptibility

Tc   T

The Ising model:
Phase transition and finite size scaling

- Is possible to determinate the critical temperature $T_{C(\infty)}$, using various lattice sizes and located the maximum of χ

$$T_{C(L)} = T_{C(\infty)} - aL^x$$

- Where $T_{C(\infty)}$ is the critical temperature of the different lattice sizes taken from the maximum of the susceptibility, a is constant to be found with the fit and x will be $-1/\nu$

- For the ferromagnetic Ising model the order parameter is the total magnetization per site.

$$M = \left\langle \left\| \frac{1}{N} \sum_{i=1}^{N} S_i \right\| \right\rangle$$

Monte Carlo method:
metropolis algorithm

The idea of the method is to find an algorithm to generate a long sequence of configurations of a system, such that after a while each configuration is generated with the adequate probability to describe the equilibrium of the system

The Metropolis Algorithm works as follow

- The system is in a state $\boldsymbol{\mu}$ and is chosen a new state $\boldsymbol{v}$

- Accepting or rejecting it at random with the chosen acceptance probability.

$$p_\mu = \frac{e^{-\beta E\mu}}{Z}$$

- If the state is accepted, the system changes to the new state v

- If not, it just leaves it as it is, and the process is repeated again and again.

Monte Carlo method:
metropolis algorithm in the Ising model

- We started with a disorder random state from the temperature after the critical point to the equilibrium temperature.

- Pick a random single spin k to be flipped.

- Calculate the difference in energy before and after the flip $E_\nu - E_\mu$.

- The change in energy between the two state is thus

$$E_\nu - E_\mu = -J \sum_{\langle ij \rangle} S_i^\nu S_j^\nu + J \sum_{\langle ij \rangle} S_i^\mu S_j^\mu$$

Monte Carlo method:
metropolis algorithm in the Ising model

- As we flip a single spin, most of the terms in the calculation in the energy different don't change and the different of energy is reduce to

$$E_\nu - E_\mu = -J \sum_{i\ n.n\ to\ k} S_i^\mu (S_k^\nu - S_k^\mu)$$

If the spin that we shoes is $S_k^\mu = +1$, then after it has been flipped, we have $S_k^\nu = -1$, then $S_k^\nu - S_k^\mu = -2$ or in the other case that $S_k^\mu = -1$ and $S_k^\nu = +1$, $S_k^\nu - S_k^\mu = +2$,

$$S_k^\nu - S_k^\mu = -2S_k^\mu$$

And so,

$$E_\nu - E_\mu = 2JS_k^\mu \sum_{i\ n.n\ to\ k} S_i^\mu$$

Monte Carlo method:
metropolis algorithm in the Ising model

- If $E_\nu - E_\mu \leq 0$ we definitely accept the move and flip the spin $S_k \to -S_k$.

- If $E_\nu - E_\mu > 0$ we still want to flip the spin with probability
  $A(\mu \to \nu) = e^{-\beta(E_\nu - E_\mu)}$.

- We choose a random number r between zero and one. If the number r is less that our acceptance ratio $r < A(\mu \to \nu)$, then we flip the spin. If it isn't, we leave the spin alone. This process is repeated over and over. Choosing spin, calculating the energy change to see if we flip it, and then deciding whether to flip it according to the acceptance ratio.

Monte Carlo method: metropolis algorithm in the Ising model
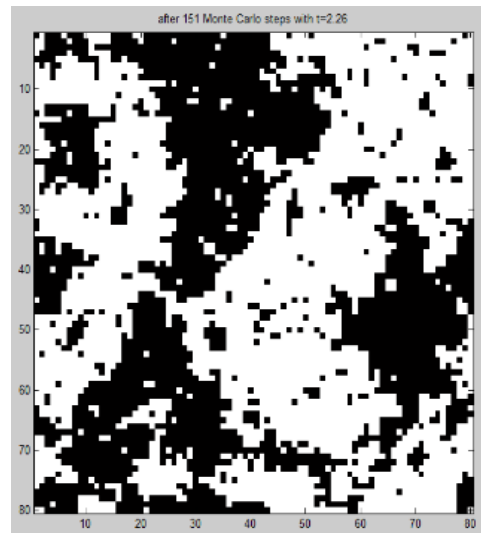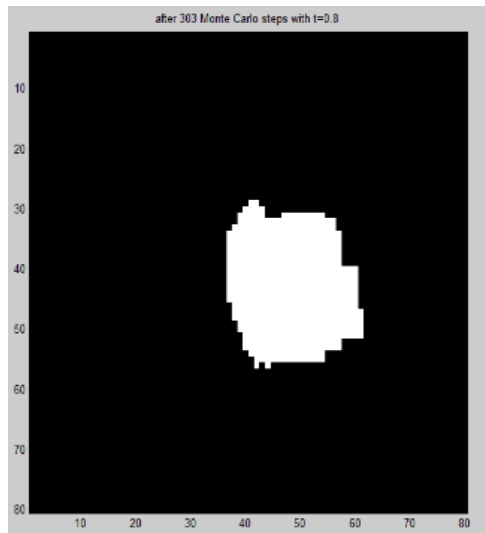
Monte Carlo simulation of ferromagnetic Ising model below Tc, over Tc, and above Tc in a finite size scale. The black dots are the spins -1 and the whites are spins +1

Monte Carlo method:
Details of Monte Carlo simulation

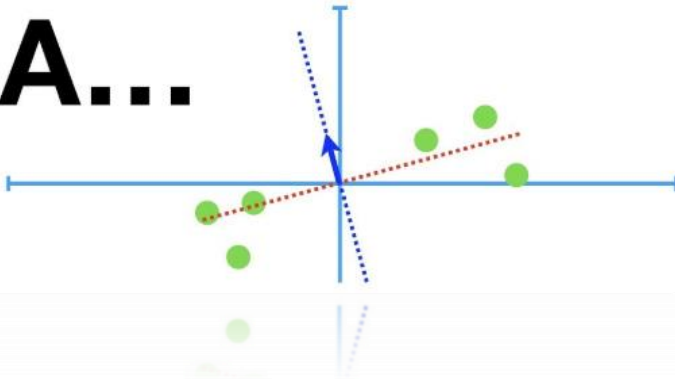| Lattice	Lattice size (L)	M.C steps	$T_{min}/$ T_{max} (T/J)	ΔT
Square	10,20,30,40,50	30000	0.8/3.5	0.025

- Ferromagnetic interaction $J = 1$

- Square system = 5000 configuration per temperature (109) having a total of configuration of 545000
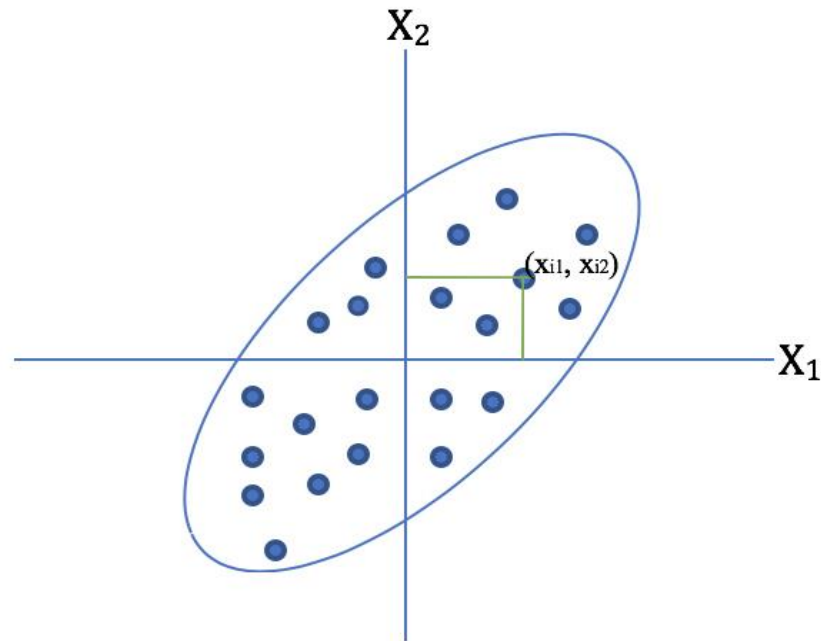
MACHINE LEARNING ALGORITHM
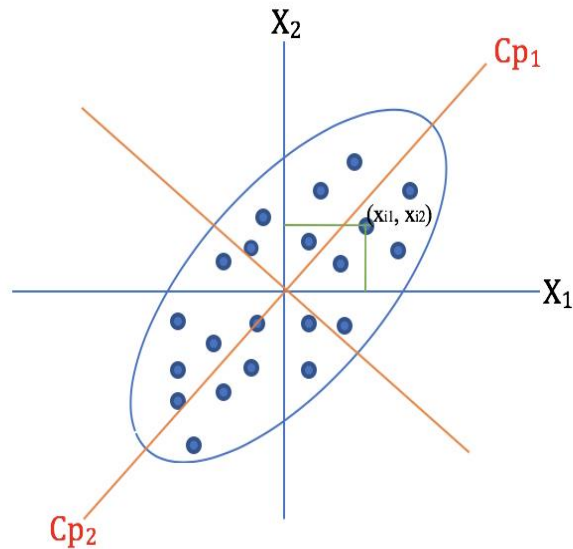
Unsupervised

MACHINE LEARNING ALGORITHMS
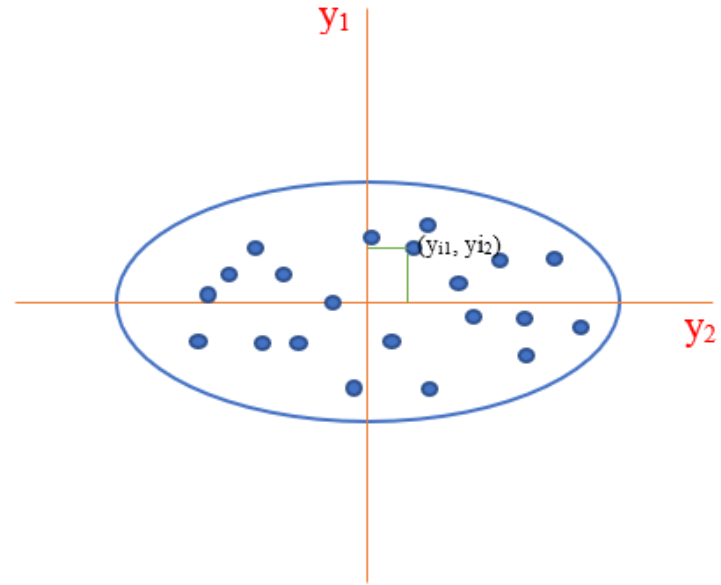**Principal Component Analysis (PCA)**

Representation of data in the variables.

MACHINE LEARNING ALGORITHMS
Principal Component Analysis (PCA)

Principal component in the data.

Principal component coordinates.

the coordinate components in terms of the initial variables are:

$$y_{i1} = x_{i1}w_{11} + x_{i2}w_{21}$$

$$y_{i2} = x_{i1}w_{12} + x_{i2}w_{22}$$

MACHINE LEARNING ALGORITHMS
Principal Component Analysis (PCA)

For all the coordinates component we have

$$
\begin{pmatrix} y_{1j} \\ . \\ . \\ . \\ y_{1N} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nN} \end{pmatrix} \begin{pmatrix} w_{1j} \\ . \\ . \\ . \\ w_{Nj} \end{pmatrix}
$$

- for all the data set of components

$$
\begin{pmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nN} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nN} \end{pmatrix} \begin{pmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{pmatrix}
$$

in matrix notation,

$$\boldsymbol{Y} = \boldsymbol{XW}$$

MACHINE LEARNING ALGORITHMS
Principal component In the Ising Model

- Data was collected by a Monte Carlo simulation in matrix X, with dimensions MxN, where $M = nT$, and N is the lattice size.

$$X = \begin{pmatrix} 1 & 1 & -1 & \cdots & -1 & 1 & -1 \\ & \vdots & & \ddots & & \vdots & \\ -1 & 1 & -1 & \cdots & 1 & 1 & -1 \end{pmatrix}_{MxN}$$

$$Y = XW$$

- The orthogonal transformation is due to vector $W = (w_1; w_2; \dots; w_N)$, where $w's$ are called weights; the first weight is found by:

$$w_1 = \arg max_{\|w\|=1} \left\{ \sum_i (x_i . w)^2 \right\}$$

MACHINE LEARNING ALGORITHMS
Principal Components in the Ising Model

- the eigenvector corresponding to the largest few eigenvalues of the matrix $X^T X (covariance)$ can be found by:

$$X^T X w_n = \lambda_n w_n$$

- The principal components are calculated as:

$$Y_n = X w_n$$

- where w_1 will be the vector corresponding to the largest variance, namely, the larger eigenvalue.

- The results are in based in the 'quantified principal components' that are defined as the average over all the configuration n at the same temperature:

$$\langle |y_n| \rangle = 1/n \sum_n |y_n|$$

Simulation details PCA
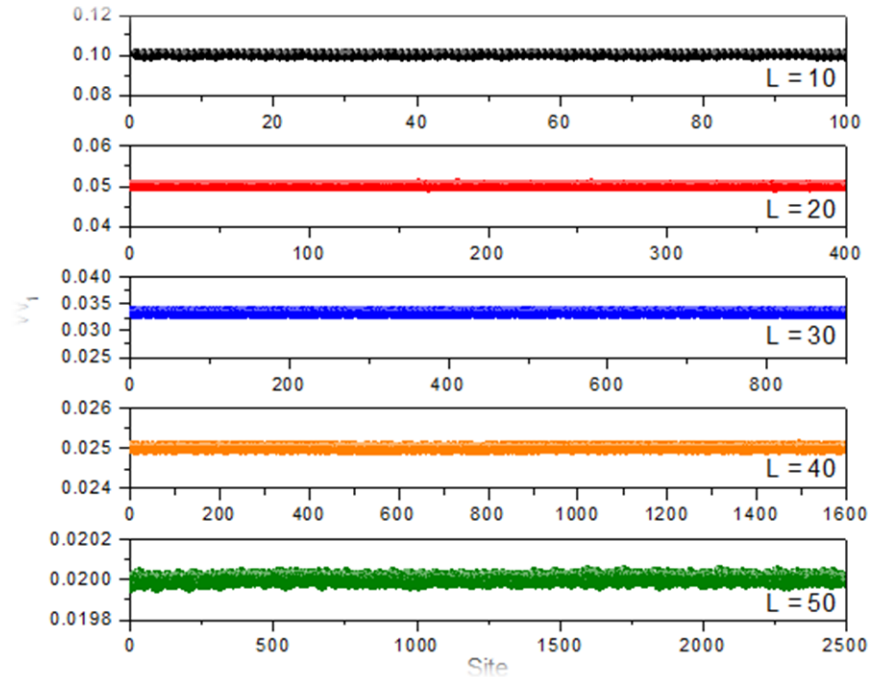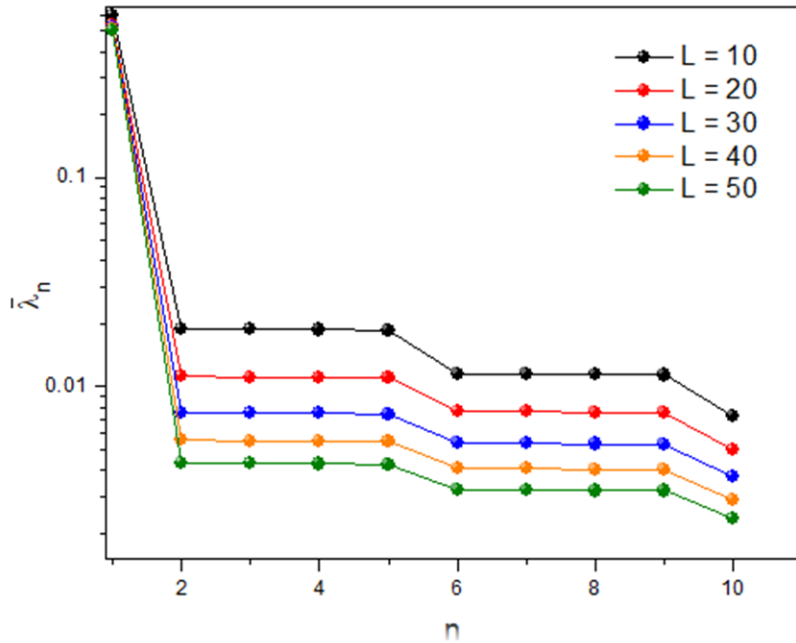
- Data from Monte Carlo Method

Square = 545000 X L

- Was performed on python

```
Xnorm = (data-np.mean(data,axis=0).T)
[latent,coeff] = np.linalg.eigh(Xnorm.T @ Xnorm)
```

Results
Square lattice PCA

PCA variance ratios from the Ising configurations for square lattice and Weights of the first principal component for each lattice size for square lattice.

Notice that $w_1 \simeq 1/L$.
Then, given a configuration from one of the rows of X.

$$Y = XW$$

$$Y_1 = \frac{1}{L}[s_1, s_2, s_3 \ldots] \begin{bmatrix} 1 \\ 1 \\ 1 \\ \ldots \end{bmatrix} = \frac{1}{L}\sum_{i=1}^{N} s_i$$

$$m = \frac{1}{N}\sum_{i=1}^{N} s_i = \frac{<Y_1>}{L}$$

Results
Square lattice PCA

If a lattice with only two spin interactions is considered and s_1 is the spin of the first lattice site, the most favorable configurations of the ordered phase of the system would be of the form:

$$x_{ord} = (s_1, s_1)$$

and for the disordered phase:

$$x_{dis1} = (s_1, s_1) \qquad x_{dis2} = (s_1, -s_1)$$

- When PCA is fed with configurations of both phases, the covariance to be computed is:

$$C = X^T X = \frac{1}{M} \sum_n x_n^T x_n$$

where, M is the total number of configurations and x_n is the nth configuration.

- If we consider a set of data p which corresponds to the ordered phase, then $(1-p)$ will correspond to the disordered phase, hence, we can write:

$$C = pC_{ord} + (1-p)C_{dis}$$

Results
Square lattice PCA

With enough data, it can be written:

$$C_{ord} = \langle x_{ord}^T x_{ord} \rangle$$

Where $< >$ denotes the average over the value that x_{ord} can take, then for this case:

$$C_{ord} = \frac{1}{2} \sum_{s_1} s_1^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
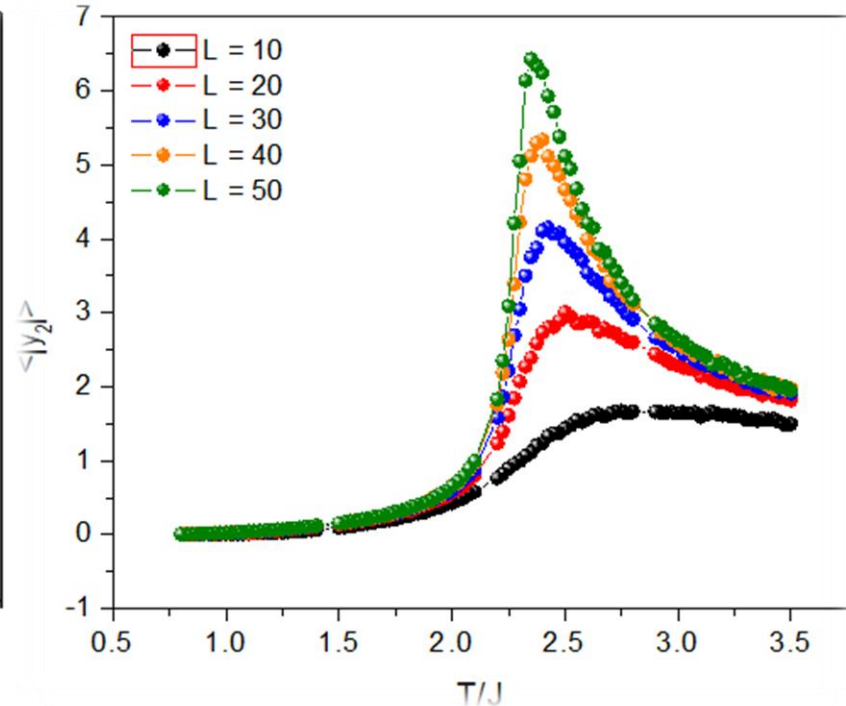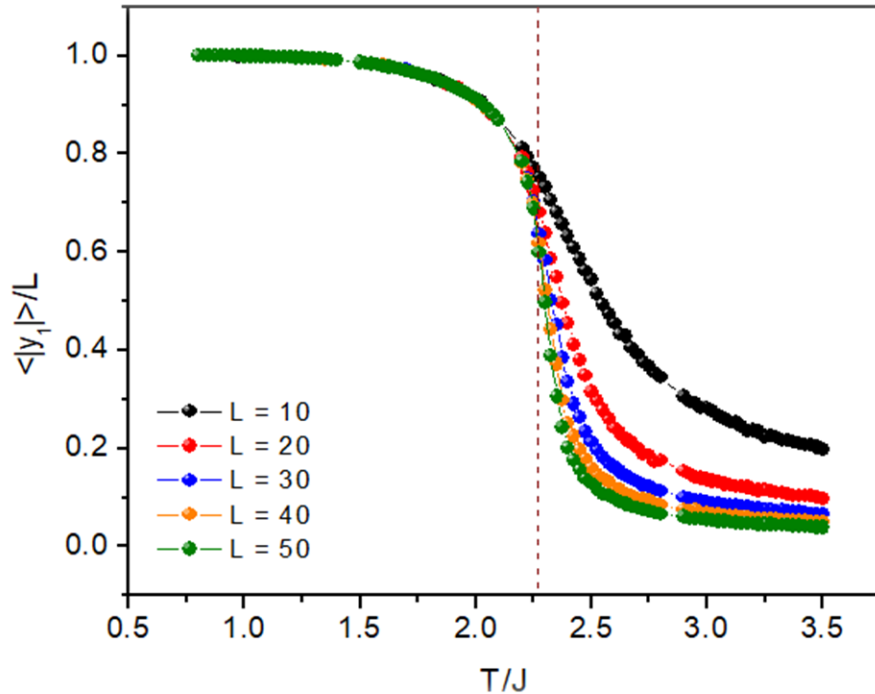
For C_{dis}:

$$C_{dis} = \frac{1}{2} \left(\langle x_{dis1}^T x_{dis1} \rangle + \langle x_{dis2}^T x_{dis2} \rangle \right) = I$$

C_{ord}, yields to the eigenvalues with its corresponding eigenvectors.

$$\lambda_1 = 2, \qquad w_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 0, \qquad w_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Results
Square lattice PCA

quantified first leading component versus temperature which represents the magnetization of the system and he quantified second leading component versus temperature which represents the susceptibility of the system

Results
Square lattice PCA

Following the in treatment based on the finite scaling relation in thermodynamic limit:
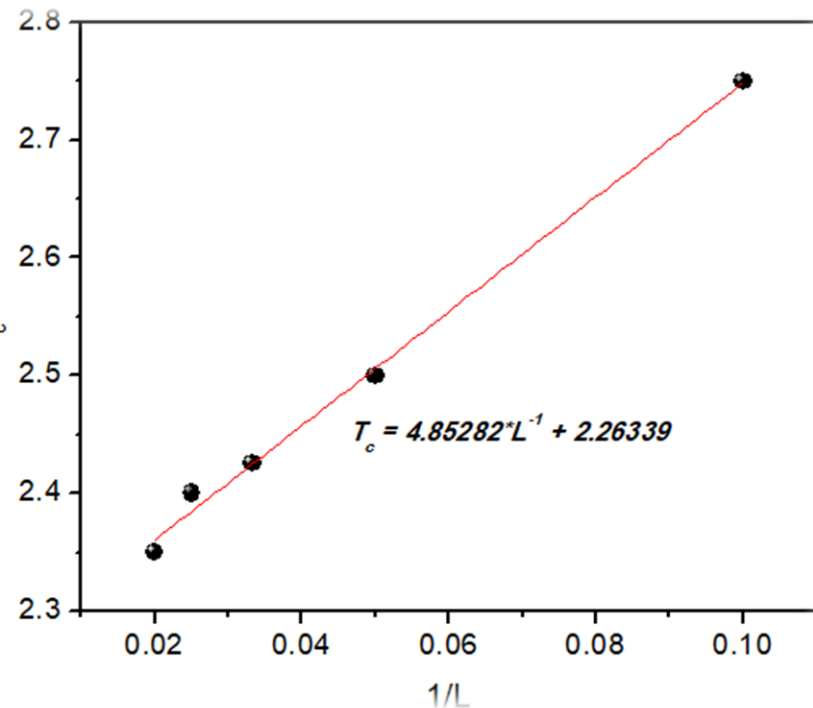
$$T_c' = T_c + KL^{-1}$$

where T_c', is the maximum of the susceptibility (χ), T_c and K are the fitting parameters.

plotting the maximums of the susceptibility graph versus the inverse of the system size, 1/L

We estimate the critical temperature $T_c = 2.26339\,J/K_B$
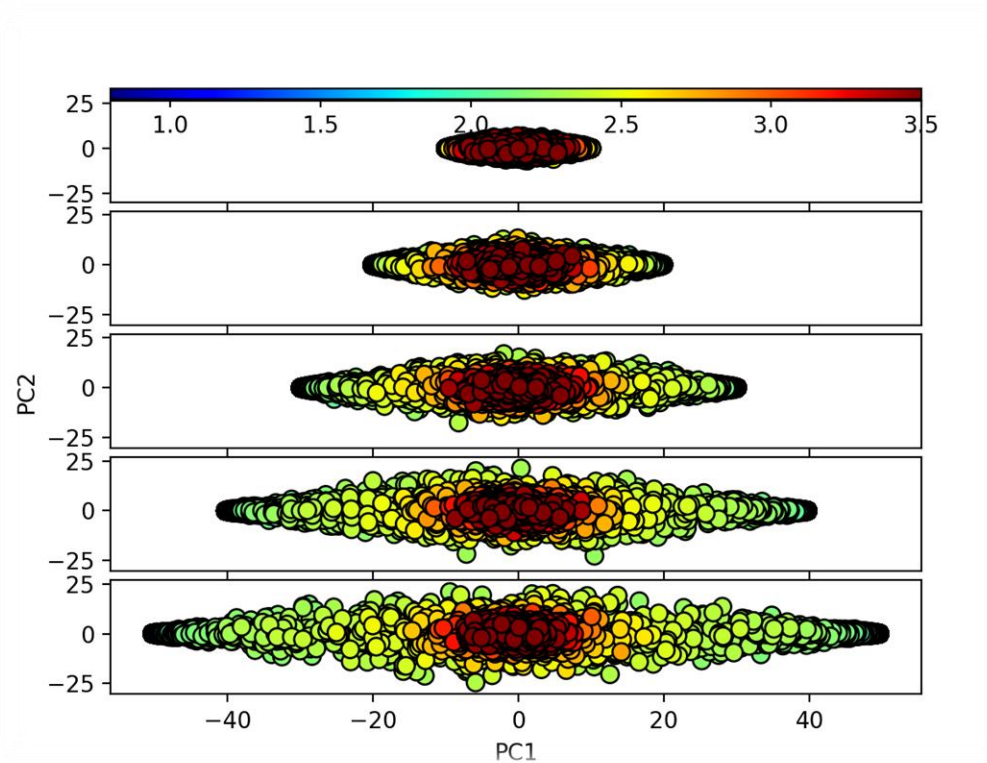
exact thermodynamic value
$$T_c = 2.2692J/K_B$$

$T_c = 4.85282{\cdot}L^{-1} + 2.26339$

Critical temperatures taken from the maximums of the susceptibility versus the inverse of the lattice size.
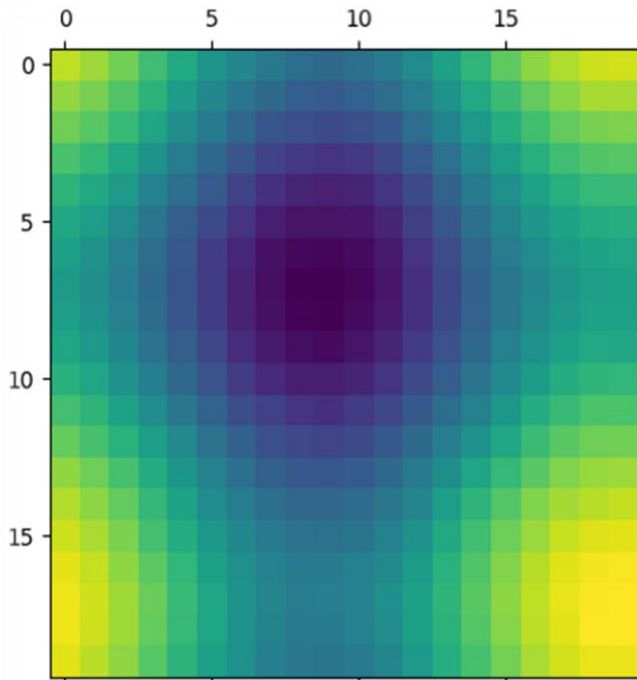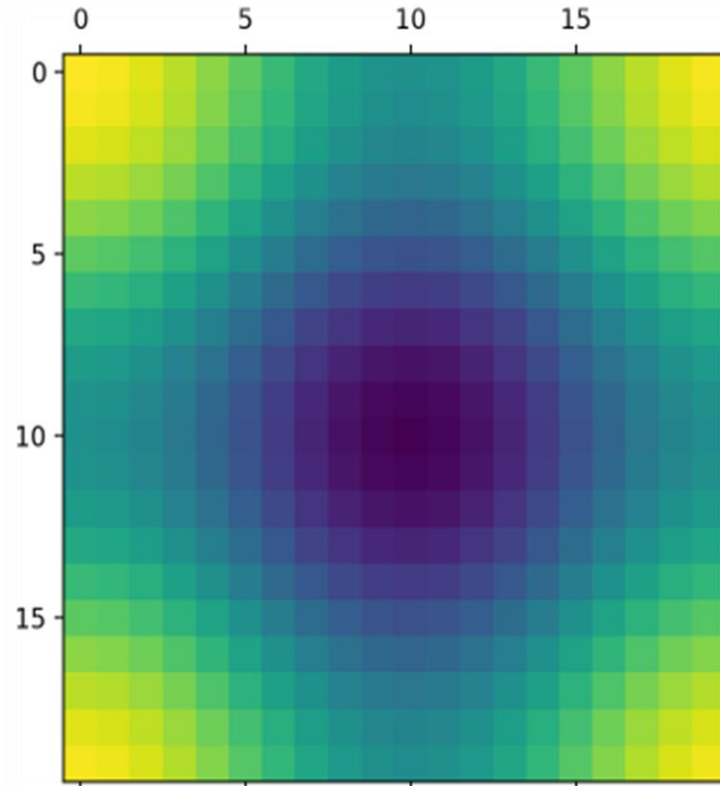
Results
Square lattice PCA

Projection of the spin configurations onto the plane for the two principal components for lattice of size 10,20,30,40 and 50 with 300 configurations for each temperature.

Results
Square lattice PCA

weight vector of the second component (w_2)

$$W_2' = \frac{1}{L}[\cos(r_1 k_1), \dots, \cos(r_N k_1)] + \frac{1}{L}[\cos(r_1 k_2), \dots, \cos(r_N k_2)]$$

where r_i is the lattice site and $k_1 = (0, 2\pi/L)$, $k_2 = (2\pi/L, 0)$ are the lowest Fourier wave vectors. The first component is associated with the origin in $k_0 = (0,0)$.

Conclusions

In this work, unsupervised machine learning technique Principal Component Analysis was employed to study phase transitions in the Ising model square system.

Principal Component Analysis was able to recognize phase transitions in the Ising model. PCA results yielded a leading component and a constant weight vector, related to this particular component. From this main component, the order parameter of the system was identified.

It was possible to mimic the magnetization and the susceptibility in the Ising model with the two-leading component. Also, the critical temperature T_c were determined. For the square system a $T_c = 2.26339 \, J/K_B$ was obtained, having a 0.5% percent error from the true thermodynamic critical temperature.

Conclusions

When PCA was feed with spin configurations from Monte Carlo, spatial order patterns were recognized by clustering the data between the ordered and disordered phases.

An interesting fact about the PCA technique for the ferromagnetic Ising model is how the weight vectors correspond to the Fourier modes of the spins configuration. In the ordered phase, the physics of the Ising model is enclosed in a single point $k_0 = (0,0)$; hence, a single dominant eigenvalue is shown.

However, careful analysis reveals that further components also encase relevant information about the system. For example, information about the susceptibility of the Ising model could be obtained from the second component. The weight vector associated with this component shows that the second Fourier mode corresponding to the spin's configurations remains as unknown that can be addressed by future research.

The others Fourier mode can potentially explain the physical meaning of other missing components, which could contribute to the Ising ferromagnetic model.

Thanks