



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2020년 2월

석사학위 논문

다층 퍼셉트론을 이용한 이차원 카고메
격자 위 이징 모형의 임계점 계산

조선대학교 대학원

물리학과

임재형

다층 퍼셉트론을 이용한 이차원 카고메 격자 위 이징 모형의 임계점 계산

Critical Point Estimation of the Two-dimensional Ising
Model on kagome lattice Using Multi-perceptron

2020년 2월 25일

조선대학교 대학원

물리학과

임재형

다층 퍼셉트론을 이용한 이차원 카고메 격자 위 이징 모형의 임계점 계산

지도교수 곽우섭

이 논문을 이학 석사학위신청 논문으로 제출함

2019년 10월

조선대학교 대학원

물리학과

임재형

임재형의 석사학위논문을 인준함

위원장	조선대학교 교수
위원	조선대학교 교수
위원	조선대학교 교수

최은서	(인)
곽우섭	(인)
박소희	(인)

2019년 11월

조선대학교 대학원

목 차

ABSTRACT

제 1 장 서론	1
제 2 장 이론적 배경	3
제 1 절 2차원 카고메 격자위에서의 이징모형	3
1. $S=1/2$ 이징 모형	3
2. 카고메 격자에서의 $S=1/2$ 이징모형	8
제 2 절 머신러닝	13
1. 다층 퍼셉트론	13
제 3 장 계산 및 분석 방법	17
제 1 절 다층 퍼셉트론의 학습	17
1. 메트로폴리스 알고리즘을 이용한 데이터의 생성	17
2. 이징모형에 대한 다층퍼셉트론의 학습방법	20
제 2 절 왕-란다우 알고리즘을 이용한 이징모형 데이터 생성	25
1. 결합 상태 밀도를 이용하는 왕-란다우 알고리즘	25
제 4 장 머신러닝, 메트로폴리스 알고리즘, 왕-란다우 방법을 이용한 데이터 비교	28
제 1 절 다층 퍼셉트론을 이용하여 생성된 데이터와 왕-란다우	

방법으로 생성된 데이터의 비교	28
1. 다층 퍼셉트론을 통해 생성된 데이터	28
2. 유한크기 축척을 통한 데이터의 비교	31
제 5 장 결 론	38

참고문헌

List of Figures

그림 1. 1차원 이징 모형의 온도 $T=0$ 에서 스핀 배열	3
그림 2. 2차원 사각 격자 이징 모형의 온도 $T=0$ 에서 스핀배열	4
그림 3. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 사각격자 위의 이징모형의 에너지	6
그림 4. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 사각격자 위의 이징모형의 비열	6
그림 5. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 사각격자 위의 이징모형의 자화도	7
그림 6. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 사각격자 위의 이징모형의 자기감수율	7
그림 7. 이차원 카고메 격자 구조	8
그림 8. 사각 격자 위 이징 모형의 스핀 배열	9
그림 9. 카고메 격자 위에서 반자성 상호작용을 하는 이징 스핀들이 격자구조로 인하여 찢절뻐뻐를 보여주는 그림	9
그림 10. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 위의 이징모형의 에너지	10

그림 11. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 위의 이징모형의 비열 11

그림 12. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 위의 이징모형의 자화도 11

그림 13. 매트릭스 폴리스 방법을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 위의 이징모형의 자기감수율 12

그림 14. 다층 퍼셉트론의 구조 13

그림 15. 은닉 층에서 사용되는 하이퍼볼릭 탄젠트 함수 14

그림 16. 정규화 선형 함수와 시그모이드 함수 15

그림 17. 온도에 따른 전이확률의 변화 19

그림 18. 다층 퍼셉트론 입력 층과 은닉 층의 연결 20

그림 19. 다층 퍼셉트론의 은닉 층과 출력 층의 연결 21

그림 20. $L=12$ 에서 강자성 이징 모형의 에너지와 자기화의 함수로서의 로그 결합 상태 밀도 27

그림 21. 다층 퍼셉트론의 오차함수 값 28

그림 22. 다층 퍼셉트론의 정확도 값 29

그림 23. 다층 퍼셉트론으로 분류시킨 카고메 격자 위의 이징 모형 자화도 30

그림 24. 다층 퍼셉트론으로 분류시킨 카고메 격자 위의 이징 모형 자기감수율 30

그림 25. 왕-란다우 표본잡기를 이용한 2차원 카고메 격자 위의 이징 모형의 자화 31

그림 26. 왕-란다우 표본잡기를 이용한 2차원 카고메 격자 위의 이징 모형의 자기 감수율
 32

그림 27. 왕-란다우 표본잡기를 이용하여 생성된 데이터의 유한 크기 축척 33

그림 28. 다층 퍼셉트론을 이용하여 생성된 데이터의 유한 크기 축척 33

ABSTRACT

Critical Point Estimation of the Two-Dimensional Ising Model on Kagome lattice using Multi-Perceptron

JaeHyoungh Lim

Advisor: Prof. Wooseop Kwak, Ph.D.

Department of Physics,

Graduate School of Chosun University

The ordered states and the disordered states of the two-dimensional Ising model on kagome lattice were studied by using multi-perceptron in length dimensions $L=4, 6, 8, 10,$ and 12 . In this study, we used the gradient descent method for minimizing the error function in the learning process of multi-perceptron. If multiple perceptron could correctly classify data at 100 percent, the accuracy is displayed as 1.

We compared and analyzed the critical temperature estimated by using finite size scaling of the peaks of magnetic susceptibilities with different linear dimensions obtained from the well-learned multi-perceptron and those obtained from the

Wang-Landau sampling method.

We had shown that the well-trained multi-perceptron could be used for estimating the critical temperature of complex spin models such as a ferromagnetic Ising model on a kagome lattice.

제 1 장 서 론

머신러닝은 기계 스스로 많은 데이터로부터 규칙성을 학습하여 기계 스스로의 기능 향상에 초점을 둔 인공지능의 한 분야이다. 1949년 Hebb의 Hebbian Learning Theory를 발표하는 것으로 시작된 머신러닝은 이후 1952년 IBM에서 근무하던 Arthur Samuel이 최초의 머신러닝 프로그램이라 말할 수 있는 체커 프로그램을 개발하였으며 머신러닝을 명확히 프로그램하지 않고도 기계에 사고하는 능력을 제공하는 것으로 정의하였다. [7]

1957년 Frank Rosenblatt는 머신러닝의 기초 알고리즘인 퍼셉트론(Perceptron)이라는 개념을 발표하였다. 이는 초기의 인공신경망(Artificial Neural Network)으로서 인간의 실제 뇌를 구성하는 신경세포의 뉴런과 동작이 비슷하다. 발표로부터 3년 뒤, Widrow, Hoff가 Delta Learning Rule을 발표하여 퍼셉트론(Perceptron)의 학습을 위한 방법으로 사용되었으나, 퍼셉트론이 XOR 문제에 적합하지 않은 문제로 인공신경망은 1980년대까지 휴먼기를 가질 수밖에 없었고, 1981년 다층 퍼셉트론(Multilayer Perceptron)과 역전파(Back Propagation)의 발견으로 인공신경망(Artificial Neural Network)의 연구가 다시 붐을 이루게 되었으며 2019년 현재까지도 활발한 연구가 이루어지고 있다. [8]

통계물리학에서는 스핀모형의 상전이(Phase Transition) 그리고 임계현상(Critical Phenomena)을 연구하기 위해서 단순한 자기 시스템(magnetic system)인 이징모형(Ising Model)을 이용하고 있다. [3, 4] 스핀-1/2 이징모형(Ising Model)에서 미시적인 자성 스핀은 +1(Upward) 또는 -1(Downward)을 가질 수 있으며, 사각 격자에서의 이징모형(Ising model)의 상전이와 임계현상의 해석적인 연구방법은 Onsager가 전단 행렬 방법(Transfer-matrix method)을 이용하여 풀었다.

이렇게 통계 물리학 및 다양한 분야에서 여러 해석적인 방법뿐만 아니라 수치 해석적으로 활발히 연구되고 있는 이징 모형(Ising Model)은 격자의 스핀들이 최 인접 스핀들과만 상호작용을 한다. 2차원 사각격자위의 이징모형(Ising model)은 상자성-강자성 상전이 현상을 이해하기 위한 단순하고 쉬운 이론을 제공하였다. 하지만 삼각격자(Triangular Lattice)나 카고메격자(Kagomé Lattice)와 같이 찢찢맹(frustration)이 존

재하는 격자에서는 외부장이 없을 때 자화도가 질서매트 변수로서의 역할을 할 수 없으므로, 임계현상을 물리적으로 이해하는데 어려움이 존재한다.

본 논문에서는 인공신경망(Artificial Neural Network)인 다층 퍼셉트론(Multi layer perceptron)을 이용하여 찢찢매(Frustrated)이 존재하는 복잡한 스핀 모형인 카고메 격자(Kagomé Lattice)위에서의 이징모형의 상전이 현상에 대하여 연구하고자한다. 컴퓨터를 이용하여 계산하는 몬테카를로 시뮬레이션(MonteCarlo Simulation)을 이용하여 만들어낸 빅 데이터를 다층 퍼셉트론에 학습시킨 후, 잘 학습된 다층 퍼셉트론을 이용하여 만들어낸 데이터가 질서 상태(ordered state)의 영역에 포함되는지 아니면 무질서 상태(disordered state)의 영역에 포함되는지를 판단하도록 하여 임계 온도(critical temperature)를 계산하고자 하며, 다층 퍼셉트론을 이용하여 얻은 결과 값을 왕-란다우 알고리즘을 이용하여 얻은 결합 상태 밀도(joint density of state)로 추정된 임계온도 값과 비교하여 찢찢매가 존재하는 시스템에서의 다층 퍼셉트론과 같은 머신러닝을 이용한 연구가 물리모형의 연구에 적절히 사용될 수 있는지에 대하여 알아보하고자 한다.

제 2 장 이론적 배경

제 1 절 2차원 카고메 격자 위에서의 이징모형

1. $S = 1/2$ 이징 모형

이징 모형은 1920년에 에른스트 이징(Ernst Ising)의 박사 학위 논문 주제로 처음 시작되었다[1]. 자성체 또는 이종합금의 간단한 모형으로 알려진 이징모형은 컴퓨터 시뮬레이션 방법인 몬테카를로 방법을 이용하여 실제로 통계 물리학의 문제를 연구에 어떻게 응용시킬 것인지에 대하여 설명하기 위하여 자주 이용된다. 이징 모형은 +1(up spin) 또는 -1 (down spin) 의 값을 가질 수 있는 스핀(spin)이라고 불리는 N개의 불연속 변수(discrete variables)의 격자를 가지며 인접한 스핀(neighboring spins)들은 교환 에너지(exchange coupling)를 통해 상호작용(interaction)한다.

1차원 이징모형은 전달행렬방법(transfer matrix method)을 통하여 쉽게 해를 구할 수 있다. 이는 $T \neq 0$ 에서는 무질서상태(disordered state)이고, 오직 $T = 0$ 에서만 질서상태(ordered state)를 보여준다. 아래, 주기적 경계조건(periodic boundary condition)을 가지고 있는 1차원 이징 모형은 2 개의 인접스핀(neighboring spins)들이 서로 상호 작용을 하고 있다.

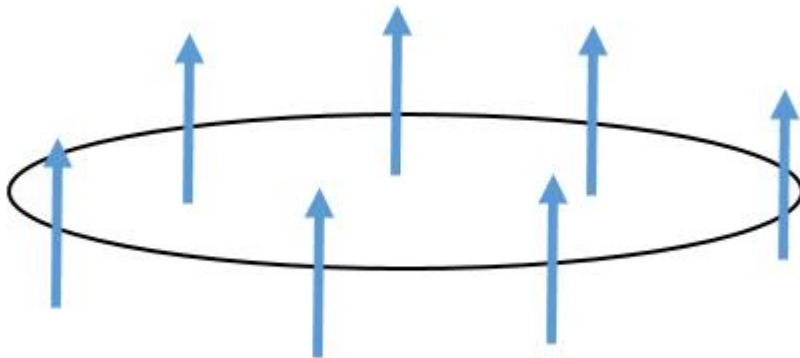


그림 2 1차원 이징 모형의 $T = 0$ 의 스핀 배열

2차원 사각 격자 위에서의 스핀들은 각각 위, 아래, 왼쪽 오른쪽에 4개의 인접 스핀들과 서로 상호작용한다.

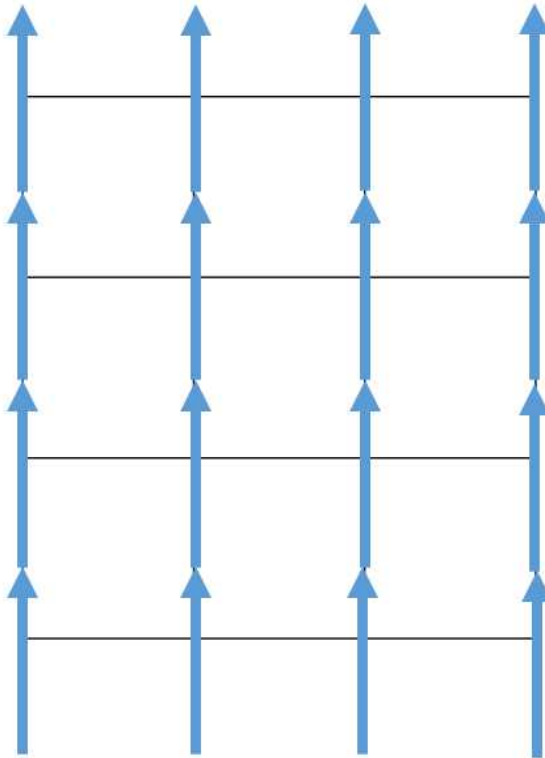


그림 3 2차원 사각 격자 이징 모형의 $T=0$ 의 스핀 배열

이 사각 격자위에 존재하는 각각의 스핀(S_i)들이 최 인접 스핀들과 결합 상수 (Coupling Constant) J 로 상호작용할 때, 스핀 모형의 총에너지인 해밀토니안 H 는 다음과 같이 주어진다.

$$H = -JE - hM \tag{2-1}$$

여기서 스핀 결합에너지(interaction energy) $E = \sum_{\langle i,j \rangle} S_i S_j$ 이며 $\langle i,j \rangle$ 는 최
 인접 쌍(nearest neighbor pair)을 나타낸다. 자화도(magnetization) $M = \sum_i S_i$
 이며 모든 스핀 값의 합을 나타낸다. 본 논문에서는 다층퍼셉트론을 기존의 전산시뮬레
 이션과 비교분석하기 위하여, 잘 알려져 있는 외부자기장(external magnetic field)
 $h=0$ 이고, 결합상수 $J > 0$ 인 강자성 모형(ferrromagnetic model)에 국한하여 논의
 한다.

2차원 사각격자에서의 이징모형은 온사거(Onsager)가 처음으로 해석적으로 정확한
 해를 구하였다. [2] 온사거가 구한 2차원 사각격자에서 이징모형의 임계온도 T_c 는

$$\sinh \left[\frac{2J}{k_b} T_c \right] = 1$$

이며, 위의 식을 정리했을 때 임계온도 $T_c = \frac{2}{\ln(1 + \sqrt{2})} \approx 2.26918531421$ 값을 얻는
 다. 즉 2차원 사각격자에서 이징모형은 $T < T_c$ 에서는 질서상태(ordered state)를
 보여주며 $T > T_c$ 이상에서는 무질서상태(disordered state)이고, $T = T_c$ 는 질서
 상태에서 무질서 상태로 전이되는 임계온도(critical temperature)이다.

온사거(Onsager)의 엄밀한 해석적 해 이외에도 컴퓨터 시뮬레이션을 이용하여 열평
 형 (Thermal equilibrium) 상태에서 열역학적 변수를 이용하여 자화도
 (Magnetization), 에너지(Energy), 자기감수율(Magnetic Susceptibility) 및 비열
 (Specific Heat) 등을 추정할 수 있으며 외부자기장이 0 일 때 열역학적 변수들
 (thermodynamic quantities)은 다음과 같다. [5]

에너지,
$$E = \left\langle -J \sum_{\langle i,j \rangle} S_i S_j \right\rangle$$

비열,
$$c_v = (\langle E^2 \rangle - \langle E \rangle^2) / NT^2$$

자화도,
$$M = \left\langle \sum_i S_i \right\rangle$$

자기감수율,
$$\chi = (\langle M^2 \rangle - \langle M \rangle^2) / NT$$

임계온도를 찾기 위해서 비열(specific heat)과 자기감수율(magnetic susceptibility)의 유한크기(finite size)에서 계산한 극값들(peaks)로 유한크기 축척(finite-size scaling)를 사용하여 2차원 사각격자위의 강자성 이징 모형에서 온사거(Onsager)와 일치하는 $T_c \approx 2.2692$ 값을 얻을 수 있다.

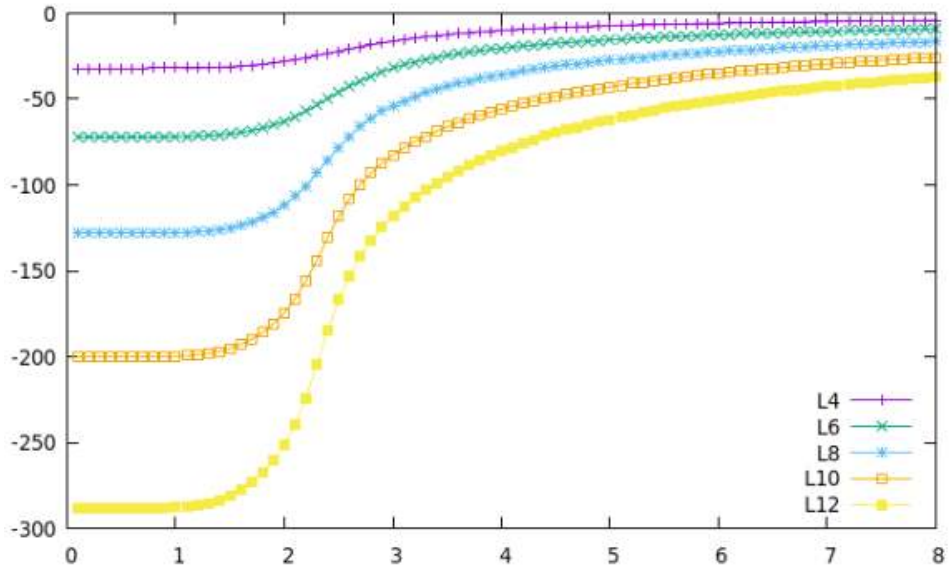


그림 4 매트릭스 폴리노미얼 전이확률을 이용한 길이차원 $L=4, 6, 8, 10, 12$ 에서의 2차원 사각격자 위의 이징모형의 에너지

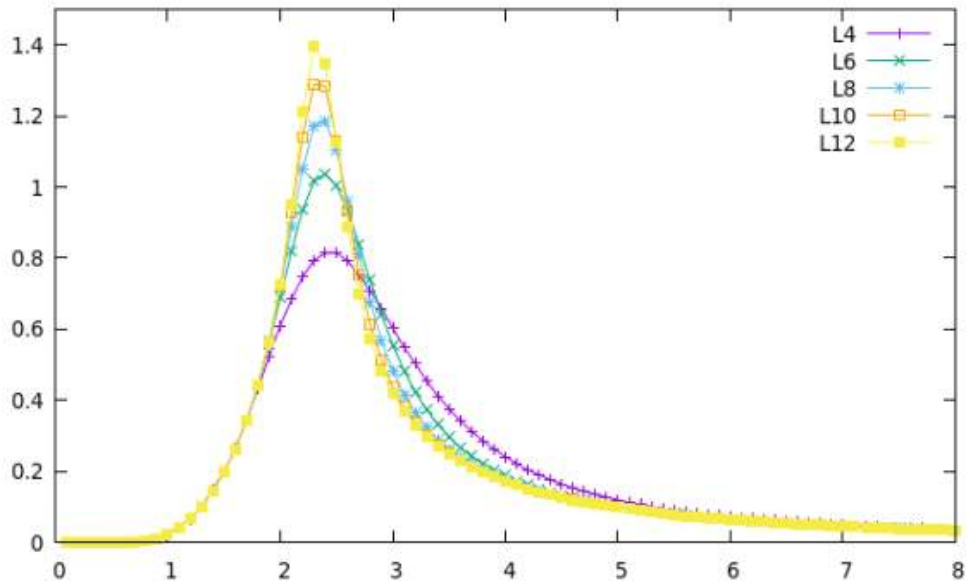


그림 5 매트릭스 폴리노미얼 전이확률을 이용한 길이차원 $L=4, 6, 8, 10, 12$ 에서의 2차원 사각격자 위의 이징모형의 비열

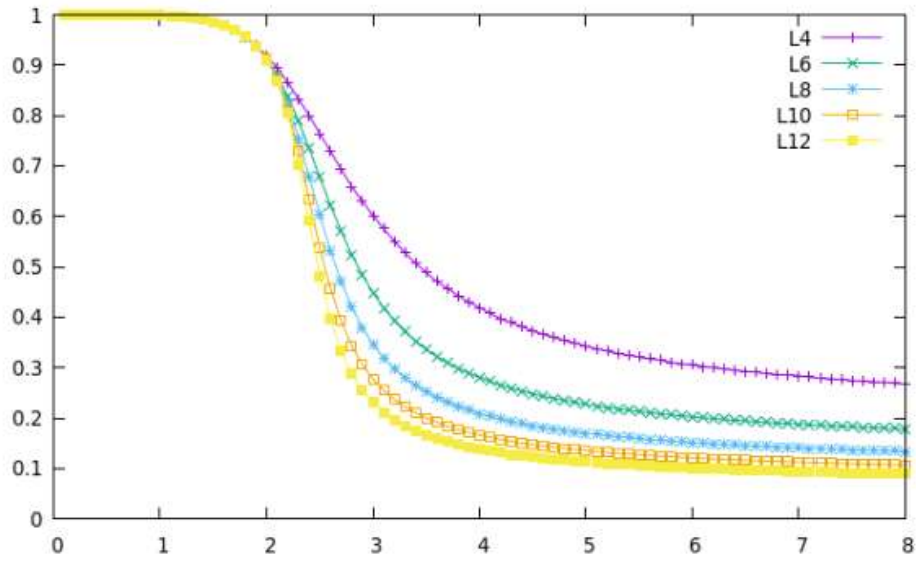


그림 6 매트릭스 폴리노미얼 전이확률을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 사각격자 위의 이징모형의 자화도

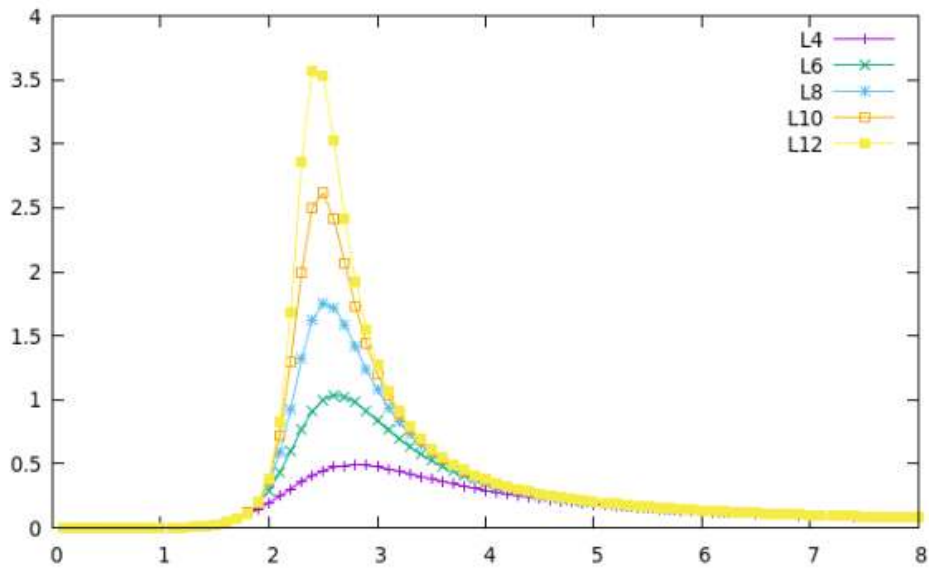


그림 7 매트릭스 폴리노미얼 전이확률을 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 사각격자 위의 이징모형의 자기감수율

2. 카고메격자 위에서의 $S = 1/2$ 이징 모형

대나무 바구니 종류중 하나인 카고메 바구니는 벌집모형 하나에 6개의 삼각형이 결합되어진 형태로 만들어져있다. [6]

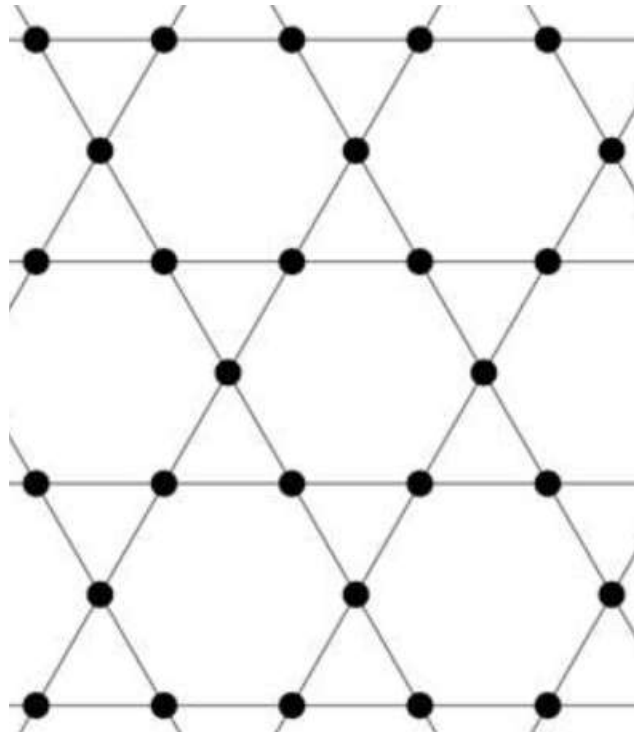


그림 8 이차원 카고메 격자구조

물리학자 쇼지(Syozji)가 재료과학(material science) 및 물리 커뮤니티(physical community)에 좌표숫자(Coordinate number)가 사각 격자(square lattice)와 같이 4 인 카고메 격자구조(kagome lattice structure)를 소개하여 널리 알려지게 되었다. [11] 카고메 격자구조의 좌표숫자는 사각 격자와 같이 4이지만 카고메 격자위의 스핀 모형의 물질의 특성은 사각 격자의 스핀모형은 같지 않다. 특히, 반자성 스핀모형이 카고메 격자 위에 있을 시에는 사각격자 위의 스핀모형과는 달리 찢찢멤 현상이 나타나고, 기저상태에서 물리적 거동 또한 지금까지 알려진 사각격자 위의 스핀모형과 전혀 다른

모습을 보여주고 있다.

반강자성 상회작용(anti-ferro magnetic interation)을 하는 사각 격자 위의 스핀 모형은 반강자성 상회작용을 하는 카고메 격자의 위의 스핀 모형과는 다르게 스핀들의 찢찢멤(frustration) 현상이 없다.

그림 8은 질서상태(ordered state)에서의 사각격자 위의 반자성 이징 모형(anti-ferromagnetic Ising model)이 스핀들이 이웃하는 스핀들과 반강자성 상호작용을 하여, 모두 다른 스핀 방향으로 에너지를 최소화 시키는 것을 보여주는 주고 있다. 중앙에 초록색 원안에 있는 스핀은 주변에 있는 최 인접 스핀들 모두와 반대 반향으로 정렬이 되어 있음을 알 수 있다.

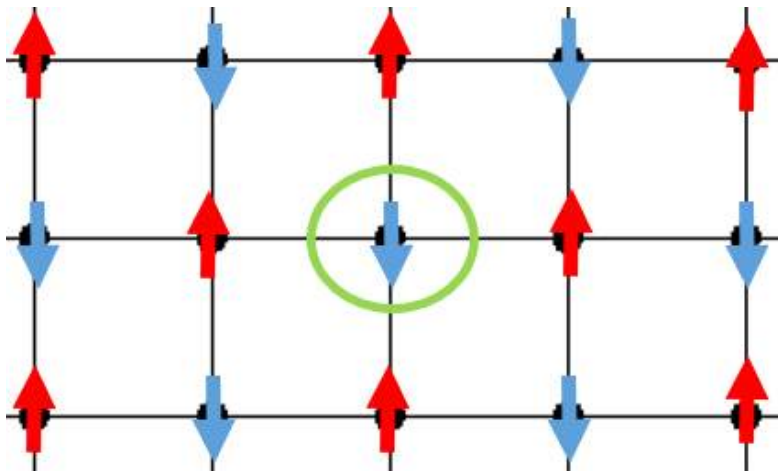


그림 9 사각 격자 위 질서상태(ordered state)에서 반자성 이징 모형(anti-ferromagnetic Ising model)의 스핀이 이웃하는 스핀들과 반강자성 상호작용을 하여, 모두 다른 스핀 방향으로 에너지를 최소화 시키는 것을 보여주는 그림

하지만, 그림 9에서 보여주고 있는 질서상태(ordered state)에서의 카고메격자 위의 반자성 이징 모형(anti-ferromagnetic Ising model)은 좌표숫자가 사각격자와 똑같은 4임에도 불구하고, 각각의 스핀들이 이웃하는 스핀들과 반강자성 상호작용을 하여, 모두 주변스핀들과 반대 반향으로 정렬되어 질 수 없는 현상인 찢찢멤이 나타나는 것을

보여주는 있다. 여기서 중앙에 있는 ? 로 표시된 위치에는 최 인접 4개의 스핀과 모두 다른 방향으로 정렬할 수 있는 스핀을 정할 수 없다. 즉, 체커보드(checker board) 모양과 같은 간단한 질서상태의 반강자성 스핀배열(spin configuration)은 존재하지 않고, 무수히 많은 기저 상태가 존재한다는 것을 알 수 있고, 이는 카고메 격자위의 이징 모형의 중요한 특징이다. 이 특징으로 인해 물리학, 재료 과학 등 여러 분야에서 찢찢매이 존재하는 시스템을 연구하기 위해 카고메 모형 위의 스핀 모형이 최근에 널리 각광을 받고 있으며 활발히 연구가 되고 있다.

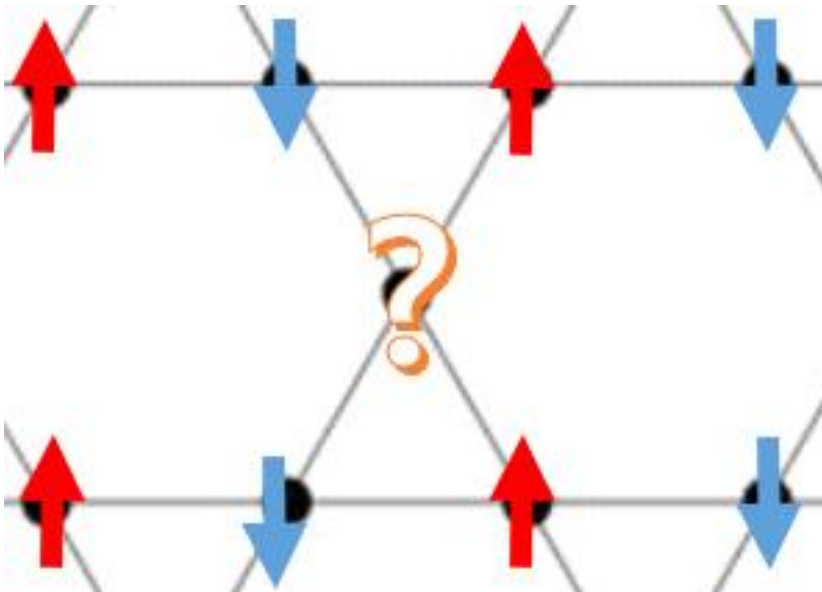


그림 10 카고메 격자 위 질서상태에서 반자성 이징 모형의 스핀이 이웃하는 스핀들과 반강자성 상호작용을 하여, 모두 다른 스핀 방향으로 에너지를 최소화 시키려고 하나, 각자 모든 스핀이 주변스핀과 반대 반향으로 정렬되어 질 수 없는 찢찢매이 나타나는 것을 보여주는 그림

이러한 카고메 격자위의 이징 모형을 컴퓨터 시뮬레이션을 이용하여 열평형 상태에서의 열역학적 변수(thermodynamic variables)인, 자화도(Magnetization), 에너지(Energy), 자기감수율(Magnetic Susceptibility) 및 비열(Specific Heat) 등을 여러 가지 시뮬레이션 방법을 사용하여 추정할 수 있다.

이징 모형의 해밀토니안은

$$H = -JE - hM$$

로 주어지고 있다. 여기서 강자성 이징 모형은 결합상수(coupling constant) $J > 0$ 이고, 반강자성 이징 모형은 $J < 0$ 이다.

널리 연구가 되어있는 외부자기장(external magnetic field) $h = 0$ 일 때, 이징 모형의 열역학적 변수들은 다음과 같이 주어진다.

에너지,
$$E = \langle -J \sum_{\langle i,j \rangle} S_i S_j \rangle / N$$

비열,
$$c_v = (\langle E^2 \rangle - \langle E \rangle^2) / NT^2$$

자화도,
$$M = \langle \sum_i S_i \rangle / N$$

자기감수율,
$$\chi = (\langle M^2 \rangle - \langle M \rangle^2) / NT$$

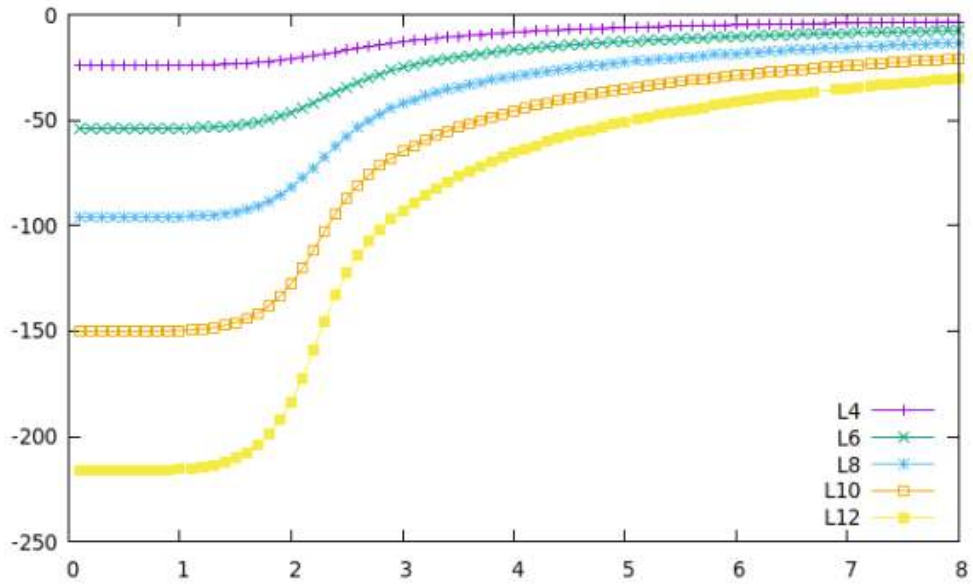


그림 11 매트릭스 폴리노미얼 전이확률을 이용한 길이차원 $L=4, 6, 8, 10, 12$ 에서의 2차원 카고메 격자 위의 이징모형의 에너지

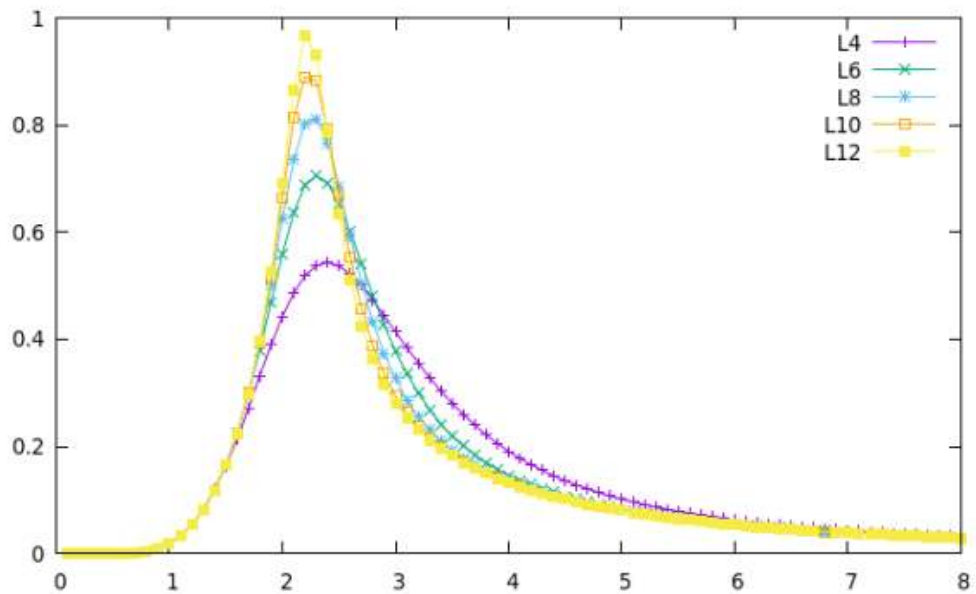


그림 12 매트릭스 폴리노미얼 전이확률을 이용한 길이차원 $L=4, 6, 8, 10, 12$ 에서의 2차원 카고메 격자 위의 이징모형의 비율

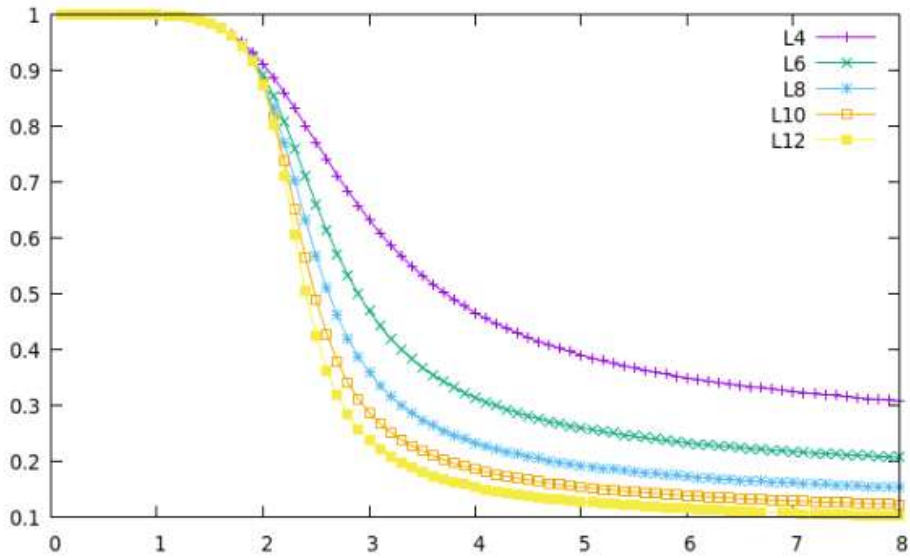


그림 13 매트로 플리스 전이확률을 이용한길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 위의 이징모형의 자화도

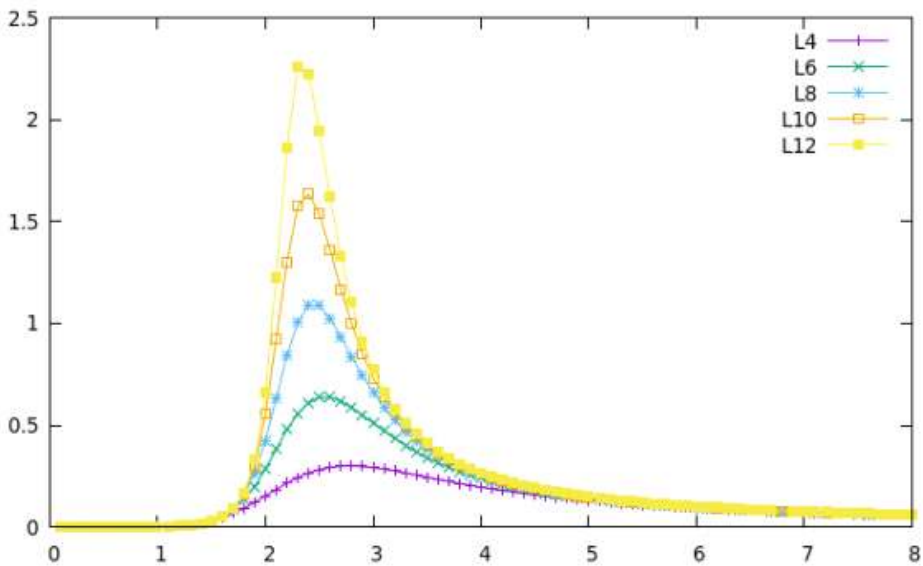


그림 14 매트로 플리스 전이확률을 이용한길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 위의 이징모형의 자기감수율

제 2 절 머신러닝(Machine Learning)

1. 다층 퍼셉트론(Multi Layer Perceptron)

1981년 시작된 다층 퍼셉트론은 인공신경망(Artificial Neural Network)으로 구성되는 학습 알고리즘의 일종이다. 이는 입력 층, 은닉 층, 출력 층 세 가지의 큰 층으로 나누어져있는 알고리즘으로 입력 층, 출력 층은 보통 한 개의 층을 가지지만 은닉 층은 학습 성능을 위하여 여러 개의 층으로 나누어져있기도 한다. [7, 8]

인공신경망의 학습방법은 대표적으로 세 가지로 나뉜다. 강화학습(Reinforcement Learning), 지도학습(Supervised Learning), 비지도학습(Unsupervised Learning), 이 세 가지의 방법은 학습 데이터의 종류와 신경망의 구성에 따라서 적절히 구분되어 사용된다. 본 논문에서는 준비된 학습 데이터인 다양한 기하학적 구조(different geometrical structure)를 가지고 있는 이차원 격자 위의 온도(temperature)에 의존하는 복잡한 이징 스핀들의 배열들을 질서적인 배열인지 또는 무질서적인 배열인지를 분류하기 위하여 다층 퍼셉트론을 사용한다. 준비한 학습 데이터로 충분히 학습되어진 다층 퍼셉트론을 통하여 분류되어진 결과 값이 질서적인 배열의 영역인지 또는 무질서적인 배열의 영역에 속하는지에 대하여 적절한 정보를 주어야하기 때문에 우리는 지도 학습을 사용한다.

여기서 우리가 사용한 다층 퍼셉트론은 데이터를 입력하는 입력 층, 데이터를 학습시키는 은닉 층, 데이터를 분류하는 출력 층 세 가지의 층으로 나누어진다.

본 논문에서의 사용하는 다층 퍼셉트론 학습 방법인 지도학습(Supervised Learning)은 출력 층에서 나온 결과 값을 미리 준비해 놓은 데이터가 질서적인 상태의 영역인지 무질서적인 상태의 영역인지를 분류한 분류 테이블인 실제 값(labeling)에 맞추어 원하는 결과가 나왔는지 분류하게 된다. 그림 14 에서 보여주는 것과 같이 $x_0 \sim x_n$ 은 입력 값으로 학습시키고 싶은 데이터를 입력시킨다. 입력되는 데이터의 값으로는 어떤 숫자든지 입력이 가능하며 입력하는 데이터의 개수도 제한은 없다. 이후 입력되어진 입력 값 데이터는 가중치(Weights) $w_0 \sim w_n$ 와 곱해진 뒤 활성화 함수를 지나 은닉 층

으로 향하게 되며 입력 층에 존재하는 1 값은 바이어스(Bias) 값으로 역시 가중치와 곱해져 활성화 함수를 지나 은닉 층으로 향하게 된다.

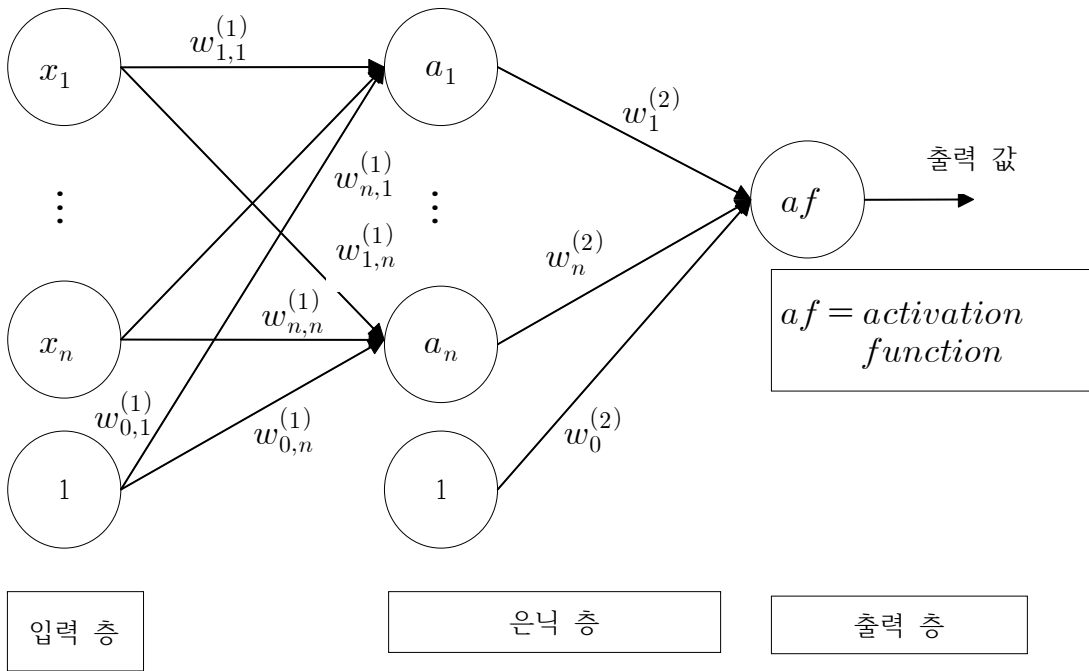


그림 15 다층 퍼셉트론의 개략적인 구조를 보여주는 그림

은닉 층에서는 a_1, a_2, \dots, a_n 의 값을 또 다시 새로운 가중치 $w_0 \sim w_n$ 와 곱하고 바이어스 값을 설정해주어 출력 층으로 그 값을 보내게 되고 출력 층에서는 활성화 함수(Activation Function)에 은닉 층에서 전달받은 값을 넣어 출력 값으로 출력한다. 은닉 층과 출력 층의 활성화함수는 같은 활성화함수를 쓰는 경우도 존재하고 다른 활성화함수를 쓰는 경우도 존재하는데, 주로 은닉 층에서의 활성화함수는 하이퍼볼릭 탄젠트(Hyperbolic Tangent) 함수를 사용하고 출력 층에서의 활성화함수는 정규화 선형 함수(ReLU, Rectified linear Unit) 또는 시그모이드(sigmoid)함수를 사용한다. 선형함수와 시그모이드 함수를 비교하였을 때, 역전파 방법(backpropagation method)에서의 학습 손실을 보정하기 위하여 미분 값이 0 인 부분이 존재하는 시그모이드 함수보다 정규화 선형 함수를 선호하나 본 논문에서는 0 이하의 입력 값들을 0으로 만드는 정규화 선형 함수보다는 0 이하의 입력 값에서도 값을 가지고 있어 질서적인 상태의 영역과

무질서적인 상태의 영역의 구분이 확실하게 가능하도록 시그모이드 함수를 사용한다.

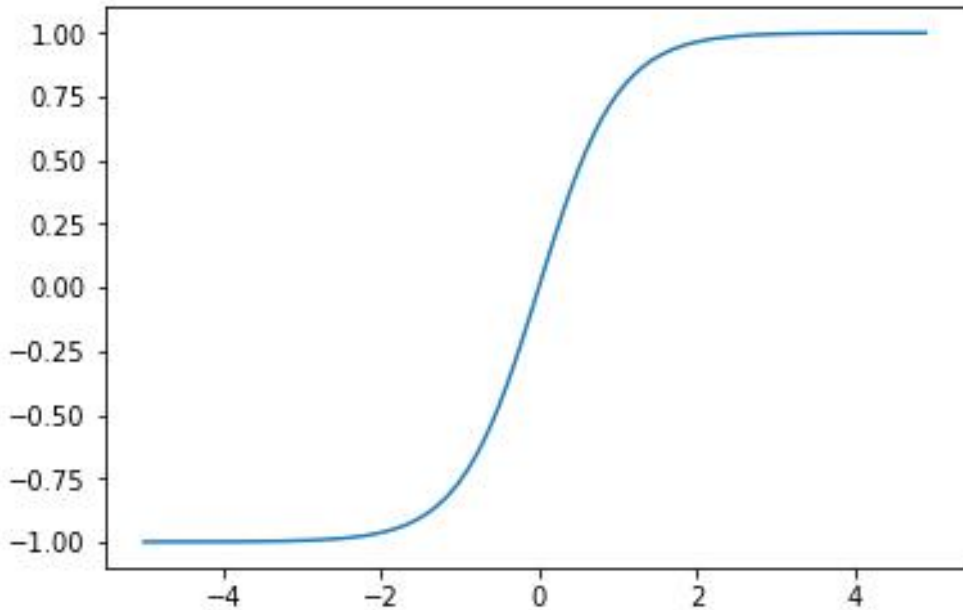


그림 16 은닉 층에서 사용되는 하이퍼볼릭 탄젠트 함수

하이퍼볼릭 탄젠트 함수는 시그모이드 함수와 비슷한 모양처럼 보이지만, 시그모이드 함수와 비교하여 $x = 0$ 에서 함수는 원점을 지나며 일정 x 값 이하 또는 이상에서의 함수 값이 1 또는 -1을 가지게 된다. 실제로 은닉 층에 정규화 선형 함수와 시그모이드 함수들이 많이 사용이 되기도 하나, 본 논문에서는 원점을 지나서 하이퍼볼릭 탄젠트 함수를 이용하여 은닉 층에서 출력 층으로 전달되는 전달 값이 입력 층에서 입력한 스핀 값들과 비슷하도록 함수 값이 1(upward) 과 -1(downward)이 되도록 만들었다.

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (\text{sigmoid})$$

$$f(x) = \max(0, x) \quad (\text{ReLU})$$

다층 퍼셉트론 출력 층에서 생성된 결과 값과 처음에 준비해놓은 실제 값(labeling)을 손실함수(Loss Function)를 이용하여 비교해 가중치를 수정한다. 대표적인 손실함수로는 MSE(Mean Squared Error)함수와 교차엔트로피(Cross entropy)함수가 존재한다.

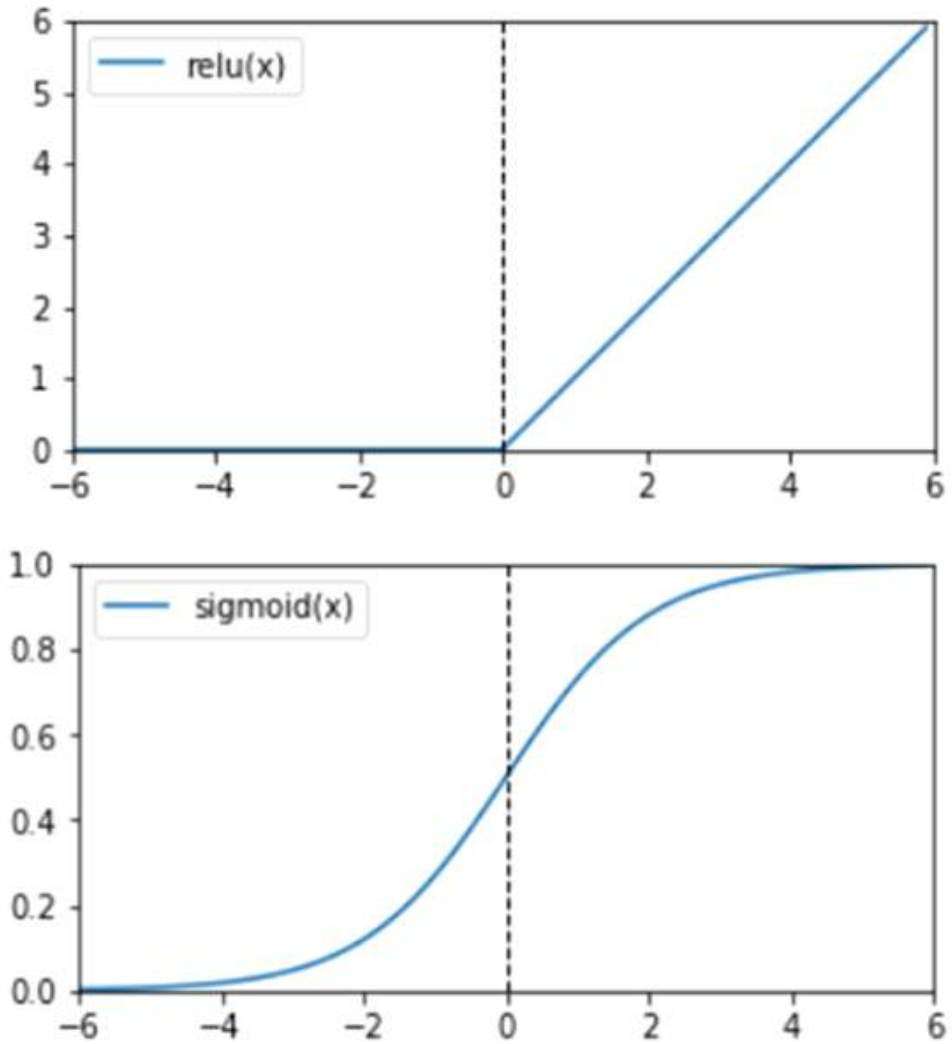


그림 17 정규화 선형 함수(위쪽)와 시그모이드 함수(아래쪽) 비교 그림

$$E = \sum \frac{1}{2} (\text{예측 값} - \text{실제 값})^2 \quad (MSE)$$

$$H(p,t) = - \sum_i t_i \ln(p_i) + (1-t_i) \ln(1-p_i) \quad (CE)$$

위에서 언급된 두 종류의 손실함수는 다음과 같은 특성을 가지고 있다. MSE 함수는 다층 퍼셉트론 출력 층에서 생성된 결과 값과 처음에 준비해놓은 실제 값(labeling)을 직접적으로 비교하여 계산하지만, 교차엔트로피함수 $H(p, t)$ 는 확률로서 다층 퍼셉트론 출력 층에서 생성된 결과 값이 처음에 준비해놓은 실제 값과 일치할 확률을 보여준다. 본 논문에서는 길이차원이 작은 격자구조에서는 MSE함수를 사용하여 직접 비교하였고 길이차원이 큰 격자구조에서는 MSE함수를 사용하기보다는 교차 엔트로피 함수를 사용하여 학습속도를 개선하고 많은 데이터에 대한 MSE함수의 단점을 보완하였다.

이렇게 MSE 함수와 교차 엔트로피 함수로 추정된 오차 E 를 최소화 시키는 방향으로 가중치 w 를 새로운 가중치 w_{new} 로 수정하기 위하여 다음과 같은 경사 하강법 (Gradient Descent)을 이용한다.

$$w_{\neq w} = w - \eta \frac{\partial E}{\partial w}$$

이 경사 하강법에서 가중치의 변화량에 따른 오차의 변화율 $\frac{\partial E}{\partial w}$ 는 역전오차파 (Backpropagation of Errors)를 사용하여 계산한다.

제 3 장 계산 및 분석 방법

제 1 절 다층 퍼셉트론의 학습

1. 메트로폴리스 전이확률을 이용한 데이터의 생성

다층 퍼셉트론을 학습시키기 위해 카고메 격자위에서의 $S = 1/2$ 이징 모형의 데이터를 메트로폴리스 전이확률을 이용하여 $T=0$ 에서 $T=8$ 까지 0.1도의 온도 간격으로 각 온도마다 약 10만개씩 준비하고자 한다. 구체적으로 메트로폴리스 전이 확률 (Metropolis Transition Rate) $W_{i \rightarrow j} = \min\left(1, e^{-\Delta E/k_B T}\right)$ 을 가지고 몬테칼로 전산시뮬을 수행하여 데이터를 생성한다.[9]

시간의 변화에 대하여 상태 i 에 있을 확률은 $P_i(t)$ 는 다음과 같이 준다.

$$\frac{\partial P_i(t)}{\partial t} = -\sum_{i \neq j} [P_i(t) W_{i \rightarrow j} - P_j(t) W_{j \rightarrow i}]$$

,여기서 $P_i(t)$ 는 시간이 t 일 때 i 상태에 있을 확률이고, $W_{i \rightarrow j}$ 는 전이 확률로서 i 상태에서 j 상태로 전이할 확률을 의미한다. 평형시스템(equilibrium system)이 열역학적 평형이 이루어 졌을 때에는 시간의 변화에 대하여 $P_i(t)$ 는 0 이므로, 세부균형 (detailed balance condtion)은 만족하는 식은 다음과 같이 쓸 수 있다.

$$W_{i \rightarrow j} P_i(t) - W_{j \rightarrow i} P_j(t) = 0$$

결국 i 상태에서 j 상태로 가는 전이 확률을 $W(i \rightarrow j)$ 과 $W(j \rightarrow i)$ 의 비는

$$\frac{W_{i \rightarrow j}}{W_{j \rightarrow i}} = \frac{P_j(t)}{P_i(t)}$$

로 적을 수 있다. 평형 상태 i 에 있을 볼츠만 요소(Boltzmann factor) $P_i = e^{-E_i/k_B T}$ 이므로, 위의 식은 다음과 같이 쓸 수 있다.

$$\frac{W_{i \rightarrow j}}{W_{j \rightarrow i}} = e^{-\Delta E/k_B T}$$

여기서, 에너지 차 $\Delta E = E_j - E_i$ 이다. 전이 확률은 유일한 것이 아니므로, $W_{i \rightarrow j} = \min(1, e^{-\Delta E/k_B T})$ 로 표현되며 위 식에서 exp함수가 전이확률을 결정지음을 보여준다. 이때 i 상태에서 j 상태로 전이될 확률의 최솟값을 0, 최댓값을 1, $\prod_{j \rightarrow i} = 1$ 로 정의하여 시뮬레이션을 진행하면 확률 $W_{i \rightarrow j}$ 메트로폴리스 전이확률 이라고 부른다. 이를 세부적으로 나누어 확인하면

$$W_{i \rightarrow j} = \begin{cases} \exp(-\beta(E(j) - E(i))) & \text{if } E(j) > E(i) \\ 1 & \text{if } E(j) \leq E(i) \end{cases}$$

로 표현 가능하며 이는 두 상태의 에너지 차에 의해서 전이 확률이 결정되며 결과적으로는 에너지가 최소화 되는 방향으로서의 전이 현상이 활발하게 이루어짐을 보여주고 있으며, 여기서 $\beta = 1/k_B T$ 이다.

본 논문에서는 2차원 카고메 격자위에서의 이징모형의 질서적인 상태를 나타내는 데이터와 무질서적인 상태를 나타내는 데이터를 이용하여 학습 시킨 다층 퍼셉트론을 이용하여 매트루 폴리스 알고리즘을 이용하여 온도별로 생성한 이징모형 데이터의 임계 온도 T_c 를 구하는 것이 목적으로 이를 위하여 간단히 질서적인 상태와 무질서적인 상태가 구분 가능한 $T = 0$ 그리고 $T = \infty$ 상태의 데이터를 학습 데이터로 사용하며, 분류 데이터로는 각 온도에서, 2차원 카고메 격자위의 이징모형 데이터를 선택하고자 하므로 케노니컬 시뮬레이션에서의 온도 값은 고정되며 지금 상태와 나중 상태의 에너지 차이가 온도와 경쟁을 하여 전이를 결정짓는다. 이때 역 온도 $\beta = \frac{1}{k_B T}$ 의 함수로

전이 확률데이터를 그려보면 그림 17 과 같은 그림이 그려진다. 이는 $T = 0$ 인 상태에서는 나중 상태의 에너지가 지금 상태의 에너지보다 큰 방향으로서는 전이가 일어나지 않음을 보여주며 그림 17에서 보이는 바와 같이 $T = \infty$ 인 상태에서는 어떠한 나중 상태도 모두 1의 전이확률을 보여주며 $0 < T < \infty$ 의 상태에서는 지금 상태의 에너지

지보다 나중 상태의 에너지가 큰 전이는 exp함수의 크기에 따라 전이 확률이 정해짐을 볼 수 있으며 $T = 0$ 에서는 $\Delta E \geq 0$ 에서는 전이 확률이 0임을 볼 수 있다.

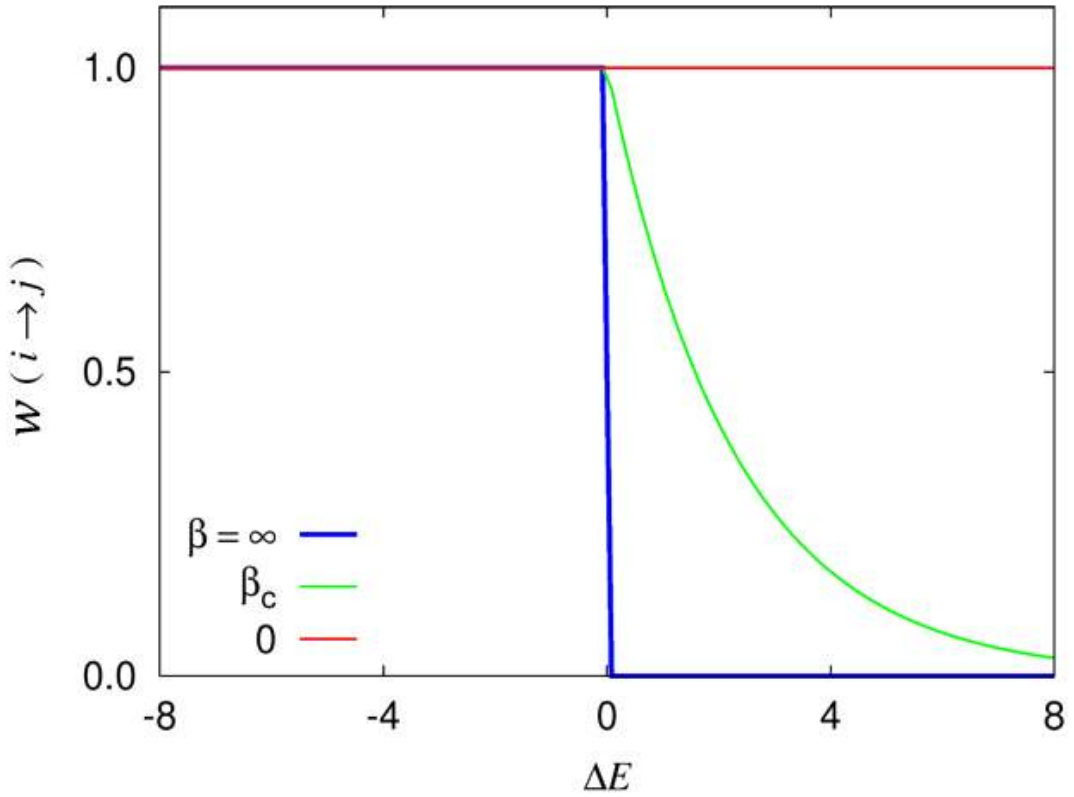


그림 18 메트로폴리스 전이확률의 온도에 따른 그래프 개형

2. 이징모형에 대한 다층퍼셉트론의 학습방법

기본적으로 2차원에서 $2n$ ($n=1,2,3 \dots$)의 길이차원을 가지는 카고메 격자 위의 이징모형의 데이터를 길이차원 $L=4,6,8,10,12$ 에서 두 온도 $T=0$ 그리고 $T=\infty$ 에 대하여 매트 로 폴리스 전이 확률을 사용하여 약 10만개의 데이터를 준비하고 데이터별로 데이터가 질서적인 상태의 영역인지 무질서적인 상태의 영역인지를 구분한 분류 테이블인 실제 값(labeling)을 준비하여 다층퍼셉트론을 학습시킨다. 매트 로 폴리스 전이 확률로 생성되는 스핀의 개수는 그 격자의 길이차원을 L 이라고 한다면 L^2 의 개수를 가지나 카고메 격자위의 이징 모형의 스핀의 개수는 사각 격자위의 이징 모형의 스핀 개수에 비하여 $3/4$ 이므로, 이에 다층 퍼셉트론의 입력 층의 노드의 개수를 $3/4 L^2$ 개로 준비하고 다층 퍼셉트론의 학습을 시작한다. [10]

1) 입력 층에서 은닉 층으로 갈 때의 학습방법

그림 18 에서 입력 층에 입력된 2차원 카고메 격자위의 이징 모형의 스핀 값들(스핀 데이터) $\{x_1, x_2, \dots, x_n\}$ 는 가중치 $w_{i,j}^{(1)}$ 가 곱해지고 입력 층에 존재하는 1에 $w_{0,n}^{(1)}$ 의 값이 곱해져 모든 값을 더해줌으로 은닉 층으로 순 입력함수 값이 정해진다. 여기서 가중치 위의 (1)은 다층 퍼셉트론의 첫 번째 연결 층이라는 의미로 본 논문에서는 은닉 층의 개수가 한 개 이므로 2번째 연결 층까지 존재한다. 허나 일반적으로는 은닉 층의 개수가 정해져 있는 게 아니므로 연결 층의 숫자는 무한히 늘어날 수 있다.

입력 층에 학습할 데이터인 2차원 카고메 격자위의 이징 모형의 각 격자의 스핀의 상태 $x_i = \{-1, 1\}$ 를 입력하면 은닉 층으로 전해지는 순 입력 함수 값을 z_i 라고 하였을 때 z_i 다음과 같이 정해진다.

$$z_j = \sum_{i=1}^n w_{i,j}^{(1)} x_i + w_{0,j}^{(1)}$$

여기서 $w_{0,i}$ 는 편향 값을 의미한다. 이렇게 구해진 순 입력 함수를 은닉 층에서 하이퍼볼릭 탄젠트 함수에 대입하게 되고 이 결과로서 나온 함수 값 a_i 이 출력 층으로 전달되게 된다.

$$a_i = \tanh(z_i)$$

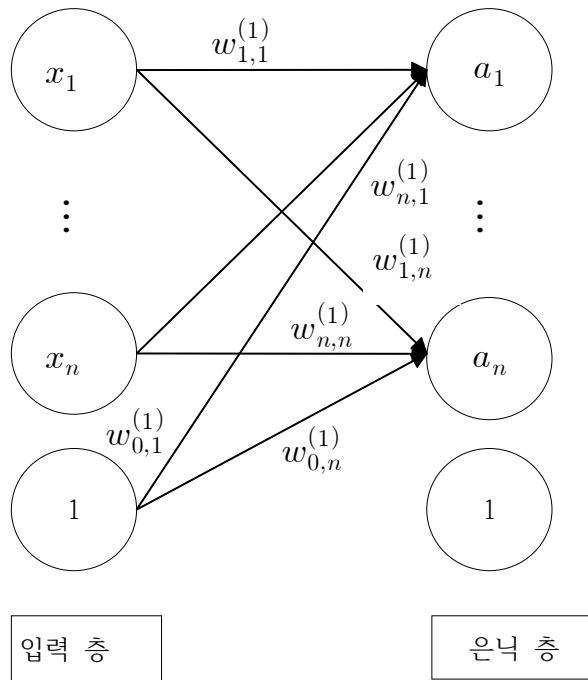


그림 19 다층 퍼셉트론의 입력 층과 은닉 층이 연결된 그림

2) 은닉 층에서 출력 층으로 갈 때의 학습방법

은닉 층에서 출력 층으로 연결은 입력 층에서 은닉 층과의 연결과 유사한 모습을 보인다. 은닉 층에서의 생성된 함수 값인 a_i 이 그림 19와 같이 출력 층으로 전해지는

데, 이 a_i 값은 입력 층에서 입력되었던 각각의 스펀 값들 $x_i = \{-1, 1\}$ 과 비슷하게 -1, 1의 값을 가지게 되며 a_i 는 새로운 가중치 $w_j^{(2)}$ 가 곱해지고, 은닉 층의 1의 값은 $w_0^{(2)}$ 가 곱해져 출력 층으로 전달되며, 이 값들을 모두 더한 함수 값은 순 입력 함수 y 로 다음과 같이 표현된다.

$$y = \sum_{i=1}^n w_i^{(2)} a_i + w_0^{(2)}$$

순 입력 함수 y 는 은닉 층과는 다르게 활성화함수인 시그모이드 함수 $\sigma(y)$ 인

$$\sigma(y) = \frac{1}{1 + e^{-y}}$$

를 통해서 결과 값 $\bar{y} = \sigma(y)$ 을 도출한다. 이후 도출된 결과 값 \bar{y} 을 다층 퍼셉트론을 학습시키기 전에 준비해 두었던 실제 값(labeling)과 비교하고 오차 함수를 구한 후, 경사하강법(gradient descent method)을 이용한 오차를 감소시키는 방향으로 학습을 진행해 나간다.

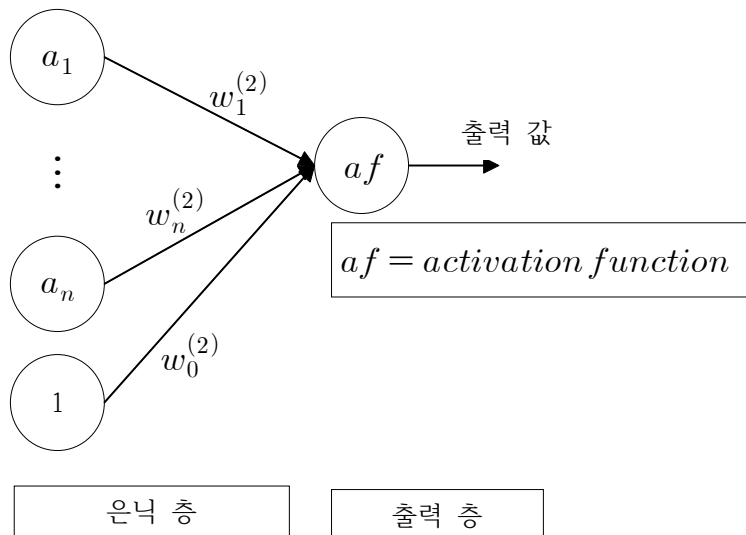


그림 20 다층 퍼셉트론의 은닉 층과 출력 층의 연결 그림

3) 오차 함수를 통한 가중치(Weights) 업데이트

다층 퍼셉트론의 학습은 입력 층 → 은닉 층 → 출력 층 순으로 진행하며, 입력된 데이터로 결과 값을 만들어 내는 순전 파(Feedforward)가 존재한다. 그리고 오차를 줄이고 가중치를 수정하기 위하여 출력 층 → 은닉 층 → 입력 층 순으로 진행되는 역전 파(Back Propagation)가 존재한다. 인공 신경망을 학습시키는 것은 최종 출력 값과 실제 값의 오차가 최소가 되도록 가중치 w 와 바이어스 w_0 를 계산하여 결정하는 일인데, 이 가중치와 바이어스를 수정하는 방법은 경사하강법으로

$$w_{\text{new}} = w - \eta \frac{\partial E}{\partial w} \quad (3-10)$$

로 나타나며 η 는 학습률을 나타낸다. 학습을 하면서 가중치에 의한 오차함수의 변화량은 노드가 늘어나면 늘어날수록 단순 수치계산으로는 계산하기가 어려워지는데 이를 보완하기 위해 역전 파를 사용하여 가중치를 수정해 나간다.

가중치를 최적화하기 위해서는 가중치의 변화에 따른 오차의 변화를 구해야한다. 수치미분으로도 구하는 게 가능하나 사실상 노드의 수가 매우 많이 늘어나면 감당할 수 있는 영역을 벗어나게 된다. 이에 순전 파를 통해서 구해진 오차를 반대 방향으로 진행하여 가중치를 수정하여 오차를 최소화 시키는 역전 파를 이용하여 가중치의 변화에 따른 오차의 변화량을 구하도록 한다. $w_1^{(2)}$ 을 수정하는 간단히 예를 들어 보면, 역전 파를 확인하기 전에 순전 파를 통해 각 노드에 전달되어지는 값들의 모습을 확인한다. 각 노드에 전해지는 순 입력함수들을 z_i 을 이용하여 표현하고 활성화함수를 a_i 을 이용하여 표현하면 입력 층에서 은닉 층으로 향하는 순전 파는

$$z_j = \sum_{i=1}^n w_{i,j}^{(1)} x_i + w_{0,j}^{(1)}$$

$$a_i = \tanh(z_i)$$

로 표현되고 은닉 층에서 출력 층으로 가는 순전 파는

$$y = \sum_{i=1}^n w_i^{(2)} a_i + w_0^{(2)}$$

$$\bar{y} = \sigma(y)$$

로 표현 가능하다. 이를 통해 가중치 w 의 변화에 대한 오차 E 의 변화 값을 미분의 연쇄법칙을 통하여 표현하면

$$\frac{\partial E}{\partial w_1^{(2)}} = \frac{\partial E}{\partial \bar{y}} \times \frac{\partial \bar{y}}{\partial y} \times \frac{\partial y}{\partial w_1^{(2)}}$$

이고 각 미분 성분을 계산하여 수식을 풀어내면 가중치의 변화에 따른 오차의 변화를 구할 수 있으므로 가중치를 경사하강법으로 업데이트 할 수 있다. 이러한 방식을 통해 다른 가중치 값들도 최적화 시킬 수 있으며 수정된 가중치 값을 가지고 다시 순전 파를 보내고 다시 역전 파를 보내어 가중치를 수정 하는 방향으로 학습을 진행한다.

가중치의 업데이트는 학습률(Learning Rate) η 값에 영향을 받는다. 학습률 η 값은 오차 함수 그래프의 기울기방향 학습에 영향을 주며 이는 최적화 과정에서의 가중치가 로컬 미니멈에 빠지지 않기 위해 적절한 값을 사용하여 주어야 한다.

제 2 절 왕-란다우 알고리즘을 이용한 이징 모형 데이터 생성

1. 결합 상태 밀도를 이용하는 왕-란다우 알고리즘

왕-란다우 알고리즘은 무작위로 스핀의 상태를 바꾸며 에너지 공간속에서의 무작위적 걷기를 시행한다. 왕-란다우 알고리즘은 각 스핀의 배열과 연관되어있는 에너지 E 와 자화도 M 의 결합 상태 밀도인 $g(E, M)$ 의 역수에 비례하는 확률로 새로운 상태를 받아들인다. 무작위적인 시행 동안 결합 공간에서 히스토그램 데이터 $H(E, M)$ 를 수집하며 각 결합 상태 (E, M) 을 방문할 때마다 히스토그램의 해당 요소를 1 만큼 증가시켜준면서, 수정 인자(Modification factor) f 를 결합 상태 밀도 $g(E, M)$ 에 곱하여 결합 상태 $A = (E, M)$ 공간에서 결합 상태 밀도를 업데이트한다.[12-15]

무작위적인 시행에서의 에너지 분포(Energy distribution)는 $g(E, M)$ 수용 확률(역수에 비례하는 확률)로 인하여 평평한 히스토그램(Flat Histogram)을 형성되어진다. 히스토그램이 평평해지면 수정인자 f 에 제곱근을 취하여 수정인자의 크기를 줄여서 시뮬레이션을 반복하게 되면 수정인자 f 는 1 에 가까워지게 되고 이때 시뮬레이션을 중단한다.

시뮬레이션을 시행하면 알 수 없는 $g(A)$ 에 대하여 추측을 하여 시작하여야 한다. 단순히 접근이 가능한 모든 상태 결합 A 에 대하여 $g(A)=1$ 로 놓으면 전체 격자에 대한 초기 배열은 마음대로 선택될 수 있다. 그 후 결합 상태 공간에서 초기 스핀 상태를 무작위로 정하고 시뮬레이션을 시작한다. 각 시행 상태는 메트로폴리스 방법과 유사하게 무작위로 스핀을 선택하여 그 스핀의 상태를 변경하여 결합상태가 A_1 에서 A_2 로 바뀌었다고 가정하면, A_1 에서 A_2 로의 전이 확률 (Transition Probability)은 다음과 같이 정해진다.

$$P(A_1 \rightarrow A_2) = \min\left(\frac{g(A_1)}{g(A_2)}, 1\right)$$

이는 $g(A_2) \leq g(A_1)$ 일 경우 무조건 A_2 가 받아들여지고 $g(A_2) \geq g(A_1)$ 의 경우에는 $g(A_1)/g(A_2)$ 의 확률로 A_2 상태를 받아들인다. 왕-란다우 방법은 메트로폴리스 방법과

비슷하게 $g(A_2) \leq g(A_1)$ 인 경우에는 무조건 적으로 상태가 받아들여지며 $g(A_2) \geq g(A_1)$ 인 경우는 0 ~ 1 사이의 무작위적인 수를 생성하여 그 수의 크기에 따라서 상태가 받아들여지는지에 대하여 결정한다. 즉, 생성된 수가 $g(A_1)/g(A_2)$ 와 같거나 작으면 상태 A_2 을 받아들인다.

만약 상태 A_2 가 받아들여지면, 보정 계수 $f > 1$ 에 따라 $g(A_2)$ 의 기존 값을 곱하여 $g(A_2) \rightarrow f \times g(A_2)$ 로 업데이트한다. 그리고 결합 상태 A_2 의 히스토그램에 기존의 요소를 업데이트한다. 만약 전이확률 상에서 A_2 상태가 받아들여지지 않고 결합상태 A_1 상태에 남게 되면 보정 계수로부터 기존의 상태 밀도 $g(A_1)$ 을 $g(A_1) \rightarrow f \times g(A_1)$ 로 업데이트한다. 그리고 결합상태 A_1 에 대한 히스토그램을 업데이트한다.[16-19]

초기 수정 인자의 최적적인 선택은 $f = f_0 = e^1 \simeq 2.71828$ 로 이루어진다. 이는 시스템에서 모든 가능한 결합상태에 빠르게 도달하도록 만들어주는 역할을 한다. 수정인자가 너무 작으면 무작위 걷기가 가능한 모든 에너지 E 와 자화 M 을 방문하는데 매우 오랜 시간이 걸린다. 허나 너무 크면 오차가 커지기에 적절한 수정인자의 선택이 중요하다.

이 무작위 시행은 히스토그램 $H(A)$ 가 평평해질 때까지 계속해서 반복하게 된다. 히스토그램이 평평해졌을 때 에너지 E 와 자화도 M 을 가지고 있는 결합상태들 A 의 방문횟수가 모두 같아지게 되면, 수정인자 f 을 줄이고 다시 무작위 시행을 진행하고를 반복하여 미리 정의된 수정인자 값(threshold) 보다 작아질 때 시뮬레이션은 끝나게 된다. 수정인자 f 는 시뮬레이션 동안 왕-란다우 방법의 정확성에 대한 제어 변수 역할을 한다.

시뮬레이션이 종료되어 생성되어진 데이터는 온도에 의존하지 않는 결합 분배함수로써, 이것을 알면 모든 열역학 량들을 계산 할 수 있게 된다.

$\text{Ln } g(E, M)$

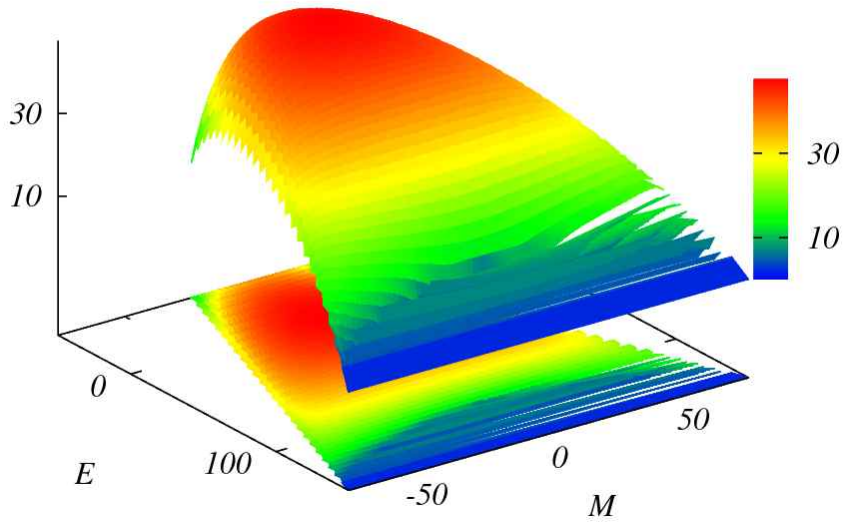


그림 21 $L=10$ 에서 카고메 격자 위에 강자성 이징 모형의 에너지와 자기화의 함수로서의 로그 결합상태밀도 $\text{Ln } g(E, M)$

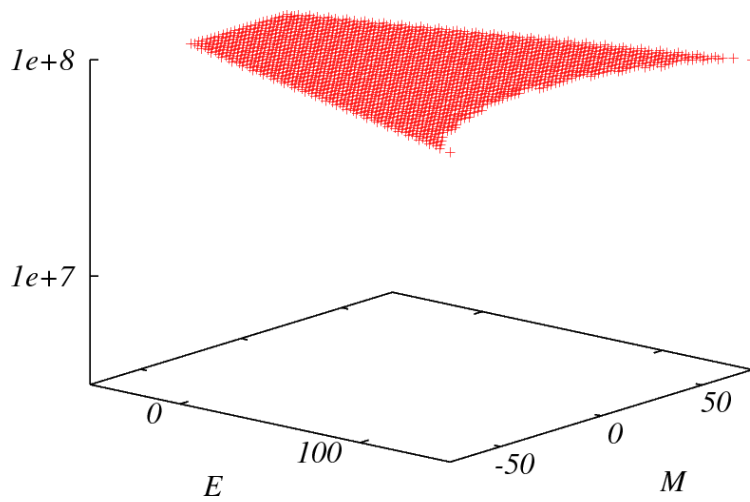


그림 22 $L=10$ 에서 카고메 격자 위의 강자성 이징 모형의 에너지와 자기화의 함수로서의 히스토그램 $H(E, M)$

제 4 장 머신러닝, 메트로폴리스 알고리즘, 왕-란다우 방법을 이용한 데이터 비교

제 1 절 다층 퍼셉트론을 이용하여 생성된 데이터와 왕-란다우 데이터의 비교

1. 다층 퍼셉트론을 통해 생성된 데이터

메트로폴리스 방법을 이용하여 온도 $T=0$ 과 $T=\infty$ 에서 2차원 카고메 격자 이징 모형의 스핀 배열 데이터를 다층 퍼셉트론에 학습시킨다. 앞서 언급한대로 본 논문에서는 다층 퍼셉트론을 지도 학습 방법에 사용하였다. 다층 퍼셉트론을 이용하여 분류되어진 데이터와 실제 값을 비교하여 올바르게 구분 되는 비율을 정확도라고 한다면 이 정확도가 1 (100%)이 되도록 다층 퍼셉트론을 학습시키고 오차 함수의 값을 0에 가깝도록 학습시킨 후 학습을 종료하였다.

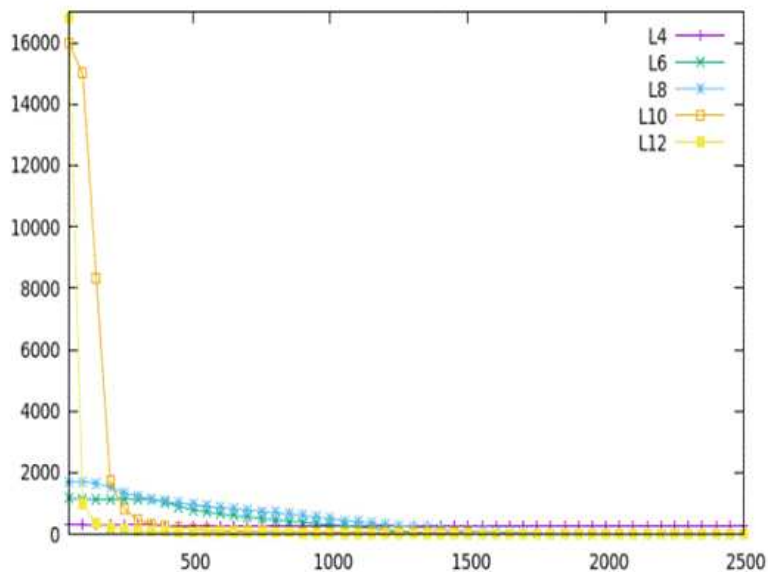


그림 23 다층 퍼셉트론을 학습시키기 위하여 정해진 오차 함수를 0에 가까운 값으로 학습시키는 모습

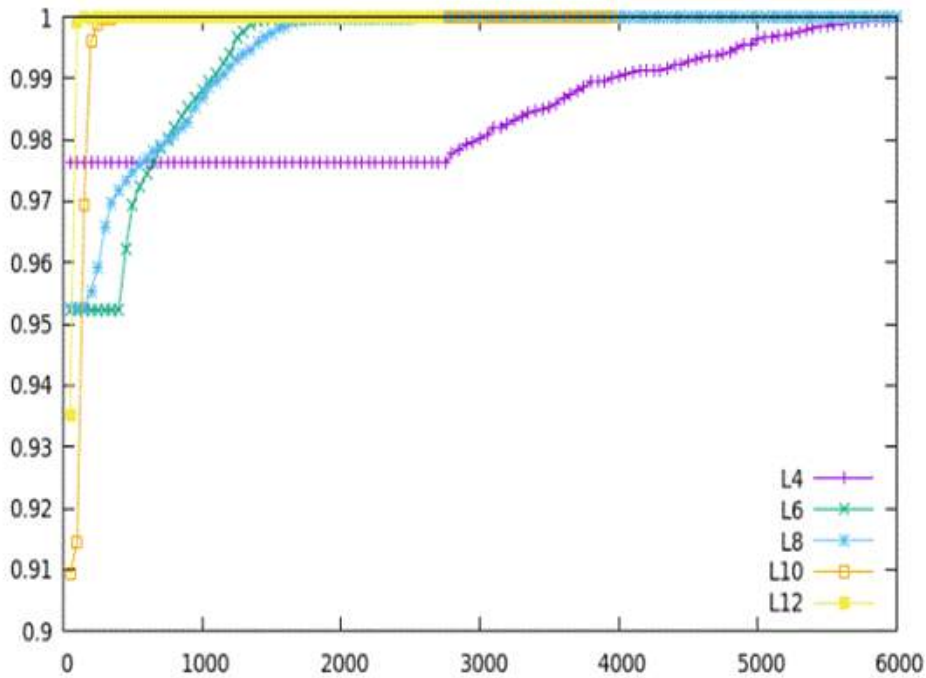


그림 24 다층 퍼셉트론의 정확도 값

그림 23 에서 이차원 카고메 격자 이징 모형을 매트릭스 풀리스 방법을 이용하여 온도 $T=0$ 에서 $T=8$ 까지 0.1도 간격으로 각 160만개씩 스핀 배열을 생성하여, 오차함수를 최소화하고 정확도를 최대화가 되도록 학습된 다층 퍼셉트론을 이용하여 데이터를 분류하였다.

그림 24 에서 시스템의 선형차원(linear dimension)의 크기가 $L=4, 6, 8, 10$ 그리고 12 인 2차원 강자성 카고메 격자 위의 이징 모형에서 메트로폴리스 방법으로 생성된 스핀 배열 데이터의 자화도를 다층 퍼셉트론으로 질서적(ordered)인지 무질서(disordered) 적인지를 분류하여 평균값을 온도의 함수로서 보여주고 있다.

그림 25 는 다층 퍼셉트론이 각각의 온도에서 분류한 메트로폴리스 스핀 배열의 자화도를 가지고 자기감수율(magnetic susceptibility)을 계산하여 온도의 함수로서 선형 차원의 크기가 $L=4, 6, 8, 10$ 그리고 12에 대하여 보여주고 있다.

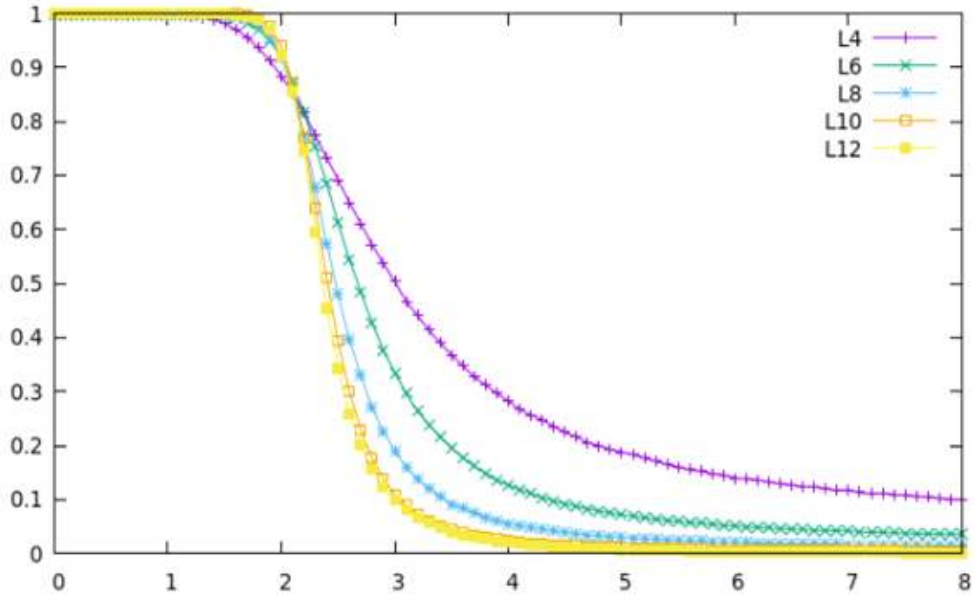


그림 25 다층 퍼셉트론으로 분류시킨 카고메 격자 이징 모형 자화

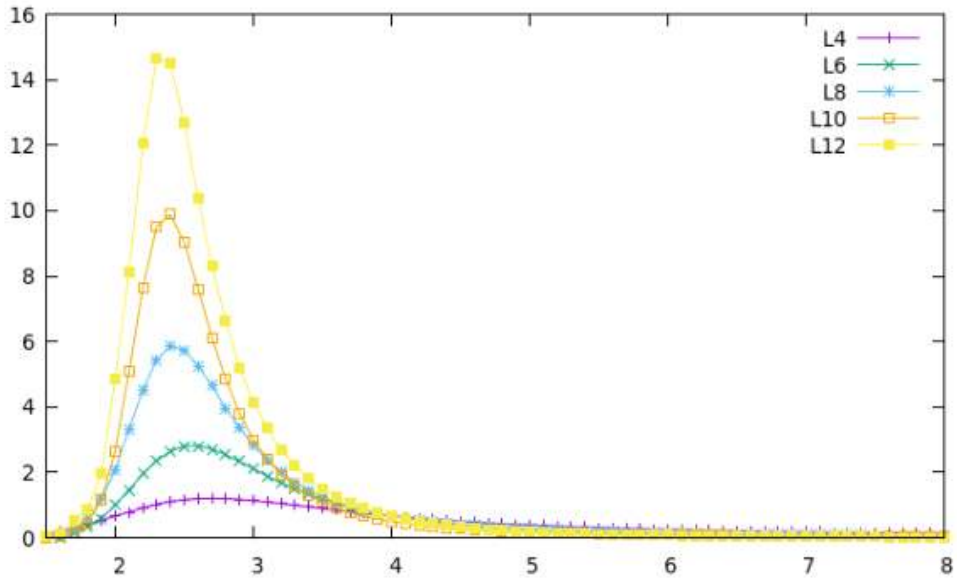


그림 26 다층 퍼셉트론으로 분류시킨 카고메 격자 이징 모형 자기감수율

2. 유한 크기 축척을 통한 데이터의 비교

왕란다우 표본잡기를 이용한 2차원 카고메 격자위의 이징모형 데이터를 이용하여 유한 크기 축척 방법을 이용하여 임계온도를 구하고자 한다. 온도의 함수로서 캐노니컬한 에너지와 자기화의 분포를 구하는 메트로폴리스 방법과 다르게 왕-란다우 표본잡기 방법은 온도와 무관한 결합 상태 밀도를 시뮬레이션을 통하여 구하는 방법으로 낮은 온도와 강한 외부장이 있어도 정확하게 열역학적 변수 등을 계산할 수 있다. 왕-란다우 표본잡기를 이용하여 길이차원 $L = 4, 6, 8, 10, 12$ 에서 생성된 이차원 카고메 격자위의 이징 모형의 데이터는 다음과 같다. 본 논문에서는 질서 매개 변수는 자화도의 절대 값으로 설정하였고, 다층퍼셉트론으로 스핀배열 데이터의 질서도를 판단하는 것이 목적이기에 에너지와 비열은 비교하지 않는다.

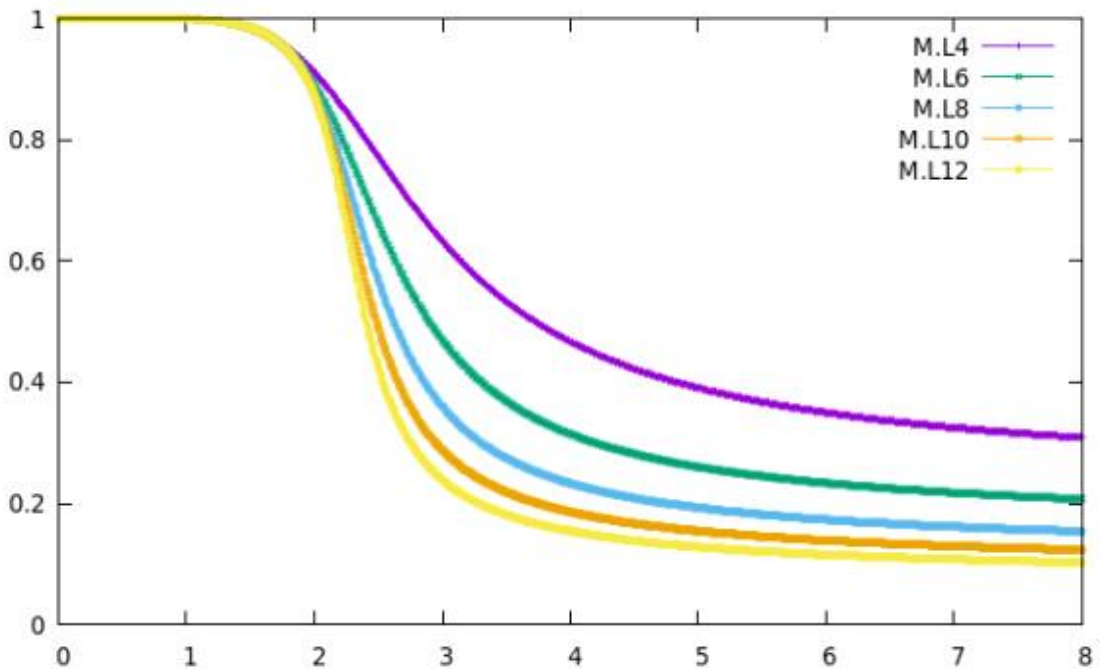


그림 27 왕-란다우 표본잡기를 이용한 길이차원 $L=4, 6, 8, 10, 12$ 에서의 2차원 카고메 격자 이징 모형의 자화도

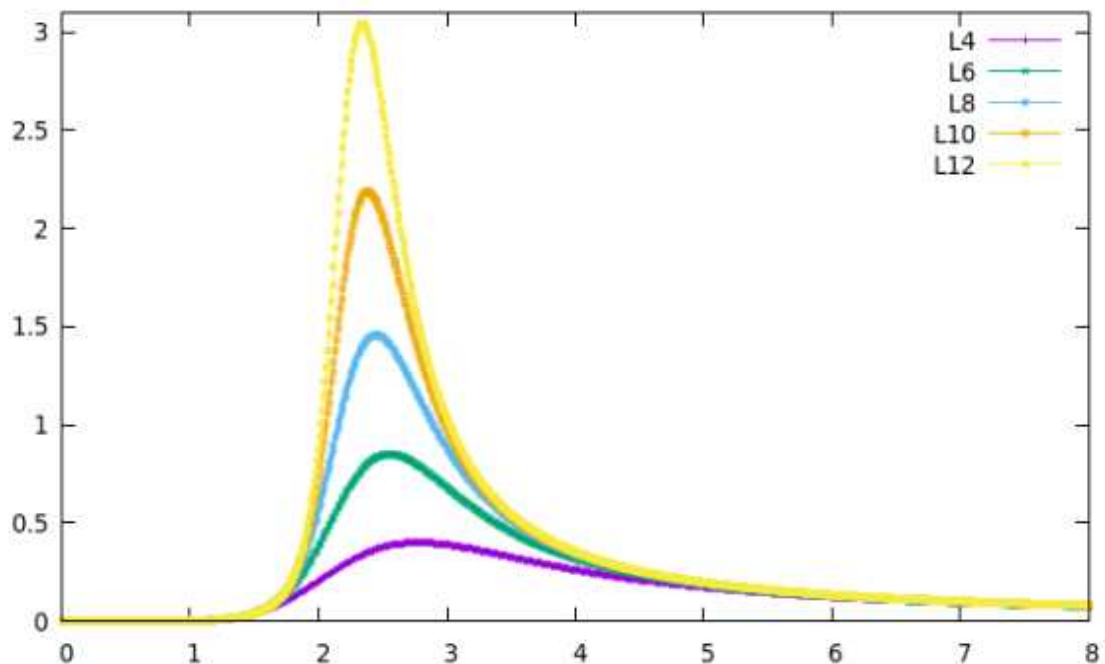


그림 28 왕-란다우 표본잡기를 이용한 길이차원 $L=4,6,8,10,12$ 에서의 2차원 카고메 격자 이징 모형의 자기감수율

왕-란다우 표본잡기를 이용하여 다양한 길이차원 L 에서 2차원 카고메 격자 이징 모형의 자기 감수율(magnetic susceptibility)의 최고점 χ_{\max} 과 유한 크기에서의 임계 온도 값 $T_c(L)$ 을 가지고 유한 크기 축척(finite size scaling)을 통하여 열역학적 극한(thermodynamic limit) $L = \infty$ 에서 임계온도 $T_c(\infty)$ 를 그림 28에서 구하였다.

그림 29는 다양한 길이차원 L 에서 메트로폴리스 스핀 배열의 자화도의 절대값인 질서매움변수를 다층 퍼셉트론으로 판단하고 각각의 유한 크기 온도 $T(L)$ 에서 계산한 자기감수율 $\chi(L)$ 의 최고점 χ_{\max} 을 가지고 유한 크기 축척을 통하여 임계온도 $T_c(\infty)$ 을 구하였다.

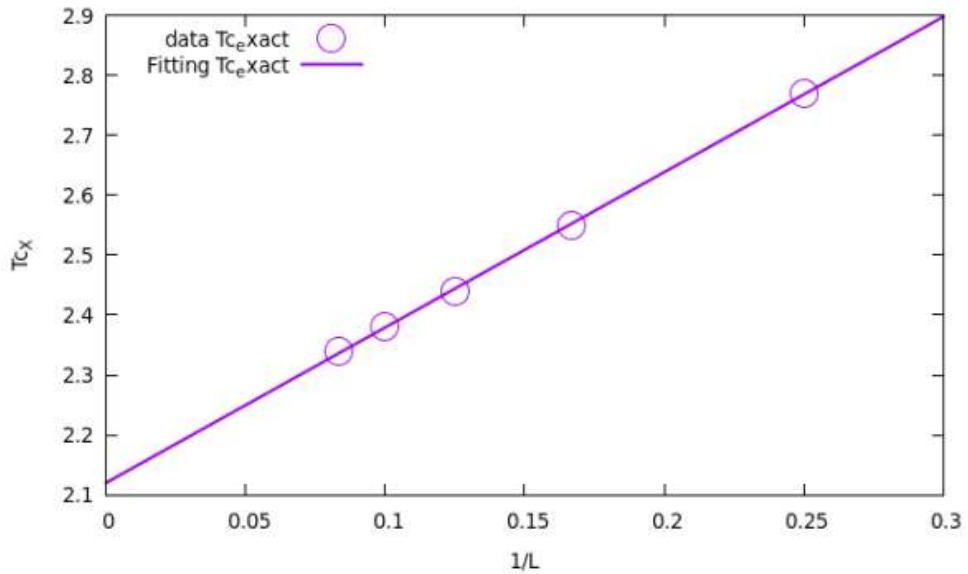


그림 29 왕 랜다우 표본잡기를 이용하여 생성된 데이터를 가지고 유한 크기 축적을 하여 임계온도를 구한 그림

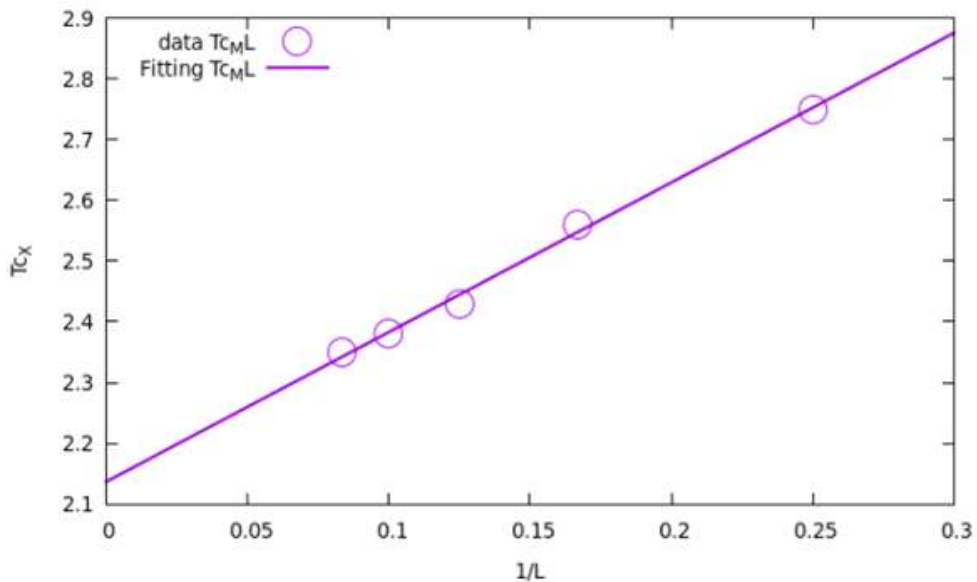


그림 30 다층 퍼셉트론을 이용하여 분류한 데이터 값을 이용하여 유한 크기 축적을 하여 임계온도를 구한 그림

두 데이터를 비교하였을 때 유한 크기 축척으로 열역학적 극한에서 왕-란다우 방법으로 계산한 임계온도가 $T_c(\infty) = 2.11(4)$ 이고 다층 퍼셉트론으로 분류하여 측정한 임계온도가 $T_c(\infty) = 2.11(5)$ 이다.

제 5 장 결 론

본 연구에서는 매트릭스 풀리스 방법을 이용하여 생성한 데이터를 다층 퍼셉트론에 학습시켜서 분류한 데이터 값을 왕-란다우 표본잡기 방법을 이용하여 생성한 데이터와 비교 연구하였다. 이차원 카고메 격자위의 강자성 이징모형의 질서매움변수(order parameter)로서 자화도(magnetization)의 절대 값을 사용하였다.

온도의 함수가 아닌 결합 상태 밀도를 계산하는 왕-란다우 방법은 시뮬레이션 시간이 매트릭스 풀리스 방법에 비하여 아주 길지만 정확한 열역학적 변수들을 계산할 수 있다.

다층 퍼셉트론으로 생성한 질서매움 변수를 이용하여 유한크기 자기감수율을 구하여 자기감수율의 최고점과 길이차원을 값을 가지고 유한크기축적으로 열역학적 극한에서 임계온도를 구하여서 왕-란다우 방법으로 계산한 값과 비교하였다. 우리는 다층 퍼셉트론을 이용하여 생성한 임계온도가 정확한 왕-란다우 방법으로 계산한 값과 유사한 값을 보임을 확인하였다. 이는 다층 퍼셉트론을 이용한 물리학적 연구의 가능성을 보여준 것이다.

본 연구에서는 인공 신경망의 가장 기초적인 신경망인 다층 퍼셉트론을 하였고, 이는 지금까지 개발되어진 다양한 인공 신경망들이 충분히 물리문제를 푸는 방법론으로 확장 가능성이 있으리라 생각된다. 앞으로 정밀한 신경망을 사용하여 데이터를 분류한다면 더욱 정확한 결과 값을 만들어 낼 수 있으리라 생각된다. 복잡한 물리문제를 풀기위하여 신경망을 이용한 데이터 연구는 더 활발히 연구되어야 할 부분이다.

참고문헌

- [1] C. Domb 1996 The Critical Point: A Historical Introduction To The Modern Theory Of Critical Phenomena (Taylor and Francis, London)
- [2] L. Onsager 1944 Phys. Rev. 65 117-149
- [3] D. P. Landau and K. Binder 2000 A Guide to Monte Carlo Methods in Statistical Physics (Cambridge: Cambridge U. Press)
- [4] D. P. Landau and R. Alben 1973 Am. J. Phys. 41 394-400
- [5] M. E. J. Newman and G. T. Barkema Monte Carlo Methods in Statistical Physics (Oxford: University press)
- [6] Seung-Yoen Kim, and Wooseop kwak “Study of the Critical Behavior of the Ising Model on a Kagome Lattice” 2018 New Phys 68 552
- [7] K. P. Murphy Machine Learning A Probabilistic Perspective (The MIT press)
- [8] Christopher M. Bishop Pattern Recognition and Machine Learning (Information Science and Statics) (New York: Springer)
- [9] Sojeong Park “Monte Carlo Simulation Studies of Critical Behavior of Spin Models” 조선대학교 대학원 이학석사학위논문 (2018)
- [10] Seung-Yoen Kim, JaeHyoun Lim, and Wooseop Kwak “Critical Point Estimation of Two-dimensional Ising model Using the Multi-Perceptron” 2019 New Phys 69 655

- [11] Seung-Yeon Kim, and Wooseop kwak “Susceptibility of the Ising Model on a Kagome Lattice by Using Wang-Landau Sampling” 2018 Korean Phys. Soc 72 653
- [12] C. Yamaguchi and Y. Okabe, J. Phys. A: Math. Gen. 34, 8781 (2001).
- [13] C. Zhou and R. N. Bhatt, Phys. Rev. E 72, 025701 (2005).
- [14] F. Wang and D. P. Landau, Phys. Rev. E 64, 056101 (2001).
- [15] T. S. Jain and J. J. de Pablo, J. Chem. Phys. 118, 4226 (2003).
- [16] Q. L. Yan, R. Faller and J. J. de Pablo, J. Chem. Phys. 116, 8745 (2002).
- [17] N. Rathore, T. A. Knotts and J. J. de Pablo, J. Chem. Phys. 118, 4285 (2003).
- [18] J.-S. Yang and W. Kwak, Comput. Phys. Comm. 179, 179 (2008).
- [19] F. Calvo, Mol. Phys. 100, 3421 (2002).

Publication

- 1) Seung-Yoen Kim, JaeHyounG Lim, and Wooseop Kwak “Critical Point Estimation of Two-dimensional Ising model Using the Multi-Perceptron”
2019 New Phys 69 655