

本科生毕业设计

|  |
| --- |
| 基于神经网络的二维Ising模型比热预测 |

|  |  |
| --- | --- |
| 院 系 | 计算机科学与技术 |
| 专业班级 | 二学位1702班 |
| 姓 名 | 刘曦 |
| 学 号 | 2017215280 |
| 指导教师 | 陈加忠 |

2020年09月16日

**学位论文原创性声明**

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。



作者签名： 2020年 9 月 16 日

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密囗，在 年解密后适用本授权书

2、不保密þ 。

（请在以上相应方框内打“√”）



作者签名： 2020年 9 月 16 日

导师签名： 年 月 日

摘 要

将机器学习运用于统计学问题的估计与预测问题在近年来得到了广泛的关注，机器学习和统计学都是描述如何从数据中学习并对其进行预测的领域，相互结合是十分便捷的，并且可以相互促进、相互学习。针对二维Ising模型与机器学习的研究比较少，针对该问题进行探讨。

从Ising模型的物理原理和研究背景开始，对模型所运用的Metropolis算法和模型所结合的反向传播神经网络架构进行了具体的描述与设计，并实现设计，计算误差，比较性能。首先描述了二维Ising模型的反向神经网络的机器学习实现过程，通过使用Metropolis算法提供该模型的期望比热及能量等参数数据，作为机器学习的训练数据集。

将数据集导入反向传播神经网络训练该算法，获得权值参数，利用该模型对任意初始状态下的二维Ising模型进行能量、比热等参数的预测，并与期望值比较。可以从结果图像中清晰地看出二维Ising模型的相变位置，并计算出相变温度，说明了该神经网络模型预测功能的有效性。最后与计算二维Ising模型的其他蒙特卡洛方法进行比较，发现Metropolis算法的模型误差更小。

**关键词**： 反向传播神经网络；二维Ising模型；Metropolis算法

Abstract

The application of machine learning to the estimation and prediction of statistical problems has attracted extensive attention in recent years. Both machine learning and statistics are fields describing how to learn from data and predict them. The combination is very convenient and can promote and learn from each other. There have been few studies on two-dimensional Ising model and machine learning, which was studied in this paper.

Based on the physical principle and research background of the Ising model, the Metropolis algorithm used in the model and the back-propagation neural network architecture combined with the model were described and designed in detail, and the design, calculation error and performance comparison were achieved. It described the realization process of machine learning based on two-dimensional Ising model's reverse neural network. Metropolis algorithm was used to provide parameter data such as expected specific heat and energy of the model as the training data set of machine learning.

The data set was imported to the back propagation neural network to train the algorithm and weight parameters were obtained. The model was used to predict the energy, specific heat and other parameters of two-dimensional Ising model in any initial state and to compare with expected value. The phase transition position of two-dimensional Ising model could be clearly seen in the result image and the phase transition temperature was calculated, demonstrating the effectiveness of the prediction function of the neural network model. Finally, it was found that the model error in this study was smaller after comparison with other Monte Carlo methods based on two-dimensional Ising model.

**Keywords:** Back propagation neural network; Two-dimensional Ising model; Metropolis algorithm

目 录

[摘 要 I](#_Toc51334392)

[Abstract II](#_Toc51334393)

[1 绪 论 1](#_Toc51334394)

[1.1 课题背景 1](#_Toc51334395)

[1.2 国内外研究现状 2](#_Toc51334396)

[1.3 课题研究目的和主要内容 3](#_Toc51334397)

[1.4 论文结构 4](#_Toc51334398)

[2 Ising模型概述 6](#_Toc51334399)

[2.1 Metropolis算法概述 6](#_Toc51334400)

[2.2 二维Ising模型概述 7](#_Toc51334401)

[2.3 预测参数的定义 8](#_Toc51334402)

[2.4 反向传播神经网络模型概述 8](#_Toc51334403)

[2.5 基于受限玻尔兹曼机的Ising模型实现 10](#_Toc51334404)

[2.6 性能比较与本章小结 10](#_Toc51334405)

[3 模型的设计 12](#_Toc51334406)

[3.1 基于Metropolis算法的二维Ising模型设计 12](#_Toc51334407)

[3.2 反向传播算法实现设计 13](#_Toc51334408)

[3.3 基于反向传播神经网络的Ising模型预测设计 13](#_Toc51334409)

[3.4 本章小结 14](#_Toc51334410)

[4 模型实现与测试 15](#_Toc51334411)

[4.1 开发环境 15](#_Toc51334412)

[4.2 模型框架实现 15](#_Toc51334413)

[4.3 模型训练实现 17](#_Toc51334414)

[4.4 模型测试实现 18](#_Toc51334415)

[4.5 结果呈现与分析 19](#_Toc51334416)

[4.6 性能对比与分析 23](#_Toc51334417)

[4.7 本章小结 24](#_Toc51334418)

[5 总结与展望 25](#_Toc51334419)

[致 谢 26](#_Toc51334420)

[参考文献 27](#_Toc51334421)

# 绪 论

本章首先介绍了本课题的研究背景以及其研究价值，接着简要分析了目前存在的统计模型预测技术所面临的一些共性问题和技术难点，以及总体上的发展趋势。随后分析了Ising模型的本质思想和该问题的普遍解决思路，简要介绍了近年来在统计学与机器学习结合方向上较为成熟的相关研究工作。最后，具体说明了本课题的组织结构与主要研究内容及工作意义。

## 课题背景

### 研究背景和意义

经典统计物理学主要研究如何使用数据来估计未知量的值，近几年开始探索生成神经网络的运用，例如受限玻尔兹曼机器[1]。许多统计物理学的中心问题都与确定和抽样状态空间的概率分布有关。因此，用生成性神经网络有效地模拟物理概率分布的可能性对于凝聚态物质、材料物理、量子计算以及涉及N体系统统计力学的其他物理科学领域具有广泛的意义[2]。由此，对于Ising模型变量组成的系统，状态空间的大小增长为，呈指数型增长，这一问题即为本论文所研讨的内容。

由于使用复杂的模型可能会导致过度拟合[3]，降低模型的预测性能。相对于复杂的偏差模型，简单的偏差模型更好，训练所需数据更少，而且由于有限大小的抽样方差而产生训练误差，会导致简单的模型在有限的抽样数下表现优于更复杂的模型。由此，本课题选择使用较为简单的反向传播神经网络模型。

对于Ising模型问题，是统计物理学中较为经典的一个模型，统计物理中引入的每一个新概念都应用于Ising模型，它最早是由物理学家威廉·楞次在1920年发明的，已有了较长的发展过程，一维Ising模型的解析解在1924年被伊辛求得[25]，二维相较于一维难很多，其解析解在1943年才被拉斯·昂萨格求出[26]。

而对于三维模型[14]目前还未找到解析解，但可用其他方法求其近似解[33]，如平均场论。

接下来对Ising模型给出定义[12]：这是一个由若干个用来描述单个原子磁矩的参数组成的系统，该参数的值代表了这些原子磁矩的自旋方向，向上或向下，即参数取1或-1，它们会受到相邻的磁矩的影响，按照某种规则排列，形成晶格。Ising模型系统可以是一维也可以是多维的，但Ising模型最高到三维，高于三维的Ising模型没有物理意义，只有抽象的数学定义，本课题讨论的二维Ising模型中每个晶格点的自旋状态都会与上下左右的四个晶格点自旋相关，该系统的可能状态共有种，计算是十分困难的。

由上文可知，研究Ising模型对于统计物理和机器学习[15]交互领域有重要意义，同时Ising模型的研究在统计物理领域也有很强的泛用性。因此进一步研究统计与机器学习的交互问题对于推动两个领域的发展有着非常重大的意义。

### 面临的问题和挑战

**模型选择的问题。**有关数据集的选取、模型的架构、成本函数的选择有许多种，因此，对于任何一个数据集的目标是获取一个对预测有效的模型，即如何选择最佳模型存在困难[3]。

**统计预测的困难。**对于统计学来说，主要研究如何使用数据来估计未知量的值，而机器学习更侧重于预测而不是估计，很好地拟合现有数据与预测新数据有根本上的不同。

**过度拟合的问题。**当训练数据量很小且数据有噪声时，会导致过度拟合，并且会大大降低模型的预测性能。

## 国内外研究现状

机器学习发端于对人工神经网络的研究，上世纪40年代以来，Warren McCulloch和Walter Pitts作为先行者提出神经网络理论，这正是现今许多流行的机器学习模型的基础[4]。1957年，康奈尔大学的Frank Rosenblatt首次用算法精确定义了自组织自学习的神经网络数据模型，设计出了第一个计算机神经网络，这是神经网络模型的开山鼻祖。1986年，Hinton等人在Nature杂志发表了著名的反向传播算法，不但明显降低了最优化问题求解的运算量，还通过增加一个隐层的方法解决了感知器无法解决的XOR Gate难题，该算法成为神经网络的最基本算法[5]。1989年，美国贝尔实验室的Yann LeCun教授提出了卷积神经网络的概念，并成功应用在了英文手写字母识别[6]。

2006年，此时计算机硬件技术的发展带来的强劲的计算能力，机器学习的泰斗Hinton提出了深度学习模型[7]。事实证明，多个隐层的神经网络具有良好的学习能力和特征提取能力。这个模型使深度机器学习进入了新时代。

2014年Baldi等人基于深度神经网络(DNN)实现了超出标准模型的新物理信号和本底判选[8]。在寻找超出标准模型的新物理中的额外重希格斯玻色子问题中，他们生成了约1100万的模拟样本，其中包含新物理中的额外重希格斯波色子信号()的样本和标准模型背景样本()。

2014年，ATLAS实验还将一个类似的数据集放到机器学习竞赛平台Kaggle上，以奖金激励的形式寻求该问题的解决方案[9]。最终，一个基于DNN集合(平均多个DNN网络输出)的方案获得了分类精度这一性能指标上的冠军。

关于Ising模型计算的发展，一维Ising模型的解析解正是由伊辛计算出来的，并且证明一维模型中没有相变，但他简单地认为二维和三维Ising模型也没有相变[25]。但如今Ising模型被认为是统计力学中基础的物理模型，正是因为它的相变过程及现象。

在1895年，Curie研究了铁磁体的铁磁-顺磁相变，所以相变点也可以称为居里点。1907年Weiss提出平均场理论，即把系统中的每个晶格点的周围的所有晶格点的作用看作一个平均的有效磁场，这虽不能计算二维Ising模型，但可以运用于三维、四维Ising模型的近似求解中[28]。在20世纪30年代，Shockley等人推进研究Ising模型中有序到无序的转变[29][30]。1941年Kramers、Wannier和Montroll精确地确定了正方Ising模型的居里点[31]。最终，在1944年，Onsager教授求解出来二维Ising模型的配分函数和比热的精确解[26]。后来，Onsager教授的一个学生Kaufman找到了一个更简便的方法，推导过程更容易理解[32]。

## 课题研究目的和主要内容

本课题的研究目的为，对于二维Ising模型，运用Metropolis算法的数值计算方法，得出一定时间过程中系统的能量、比热、磁化率等参数的变化数据。将这些数据导入神经网络中进行学习，使该程序得到与数值计算数据相同的趋势情形，即可运用该程序预测一般情形下二维Ising模型的比热等参数。

本课题的研究主要分为以下几个环节进行展开：

1. 对二维Ising模型进行理论推导与计算，运用数值模拟计算该模型在一定条件下随时间的能量等参数的变化数据。
2. 探究反向传播算法的实现原理，对其理论上的可行性和正确性做出相应的推演。随后基于反向传播神经网络模型对Ising模型进行学习，测试机器学习内容的正确性。
3. 生成随机初始状态对该网络进行功能测试与性能测试，将预测的能量、比热等参数变化趋势与数值计算的能量、比热等参数随温度变化趋势进行对比。

## 论文结构

本课题的主要内容如下：

第一章首先介绍了Ising模型的研究背景，并对Ising模型给出了详细的定义，阐明了该模型的物理思想，分析了其研究价值。接着简要阐述了目前二维Ising模型状态预测所面临的一些共性问题和技术难点，以及总体上的发展趋势。简要说明了选择反向传播算法的理由。随后论述目前的神经网络问题发展现状，介绍国内外在神经网络领域的相关研究工作，以及Ising模型计算的发展历史，最后对本课题的主要研究内容与其意义作了具体说明。

第二章首先定量描述了Ising模型计算所运用的Metropolis算法，以及其用于数值模拟Ising模型的演算过程和算法的基本流程。接着给出了本课题需要预测参数的定义式。最后介绍本课题所运用的反向传播神经网络模型的基本思想与实现过程。

第三章对于模型进行设计的说明，重点阐述了反向传播神经网络的设计实现，用矩阵的方式直观地描述其过程，并通过流程图说明了基于bp神经网络的二维Ising模型的设计流程。

第四章详细给出了基于反向传播神经网络的二维Ising模型的初始架构，对模型的网络架构、训练框架、测试框架进行说明阐述，并导入学习数据进行学习，接着测试随机数据，对模型的可行性进行分析，并与其他蒙特卡洛方法进行性能比较。

第五章中总结性地对本次毕业设计所完成的工作进行了阐述，接着对本项目可能的后续工作进行了计划，最后展望了统计学问题与机器学习今后可能的发展方向。

# Ising模型概述

在编写实现机器学习的程序之前，必须要充分了解二维Ising模型的理论知识与数学计算。为了更好地理解深度学习是如何应用于Ising模型的，以及更全面地认识统计学与机器学习领域相结合的发展现状与方向，本章主要介绍Ising模型所运用的Metropolis算法，以及相应的演算实现过程。在本章后半段中，重点论述了本课题所使用的网络模型，反向传播神经网络模型的原理与实现流程。

## Metropolis算法概述

在第一章中，简要说明了Ising模型的定义，但仅仅是定性地分析了Ising模型的原理背景与发展现状。本课题需要得到该模型的比热等参数，需要用到数值计算方法。在数值分析算法中，Metropolis算法[10]是过去蒙特卡洛方法的所有计算方式中最成功和最具影响力的。当在没有精确的分析方法，甚至没有线性算法的情况下，使用该算法，原理是从某一概率分布中抽取随机样本序列，通过这些序列可用来估计其概率分布或期望值等。

Metropolis算法的基本实现过程[11]为：

1. 初始化：选定初始状态，令时间t=0；
2. 迭代：从某一容易抽样的分布中随机生成状态x，计算是否接受候选状态的概率，接着随机生成从[0,1]均匀分布的随机数，若该随机数不大于候选状态的概率，则接受该状态并随机生成候选状态，否则放弃该状态。并令t=t+1，不断迭代直到某些条件达成或达到迭代次数为止。

基于上述原理，图 2‑1给出了Metropolis算法实现过程的流程图：

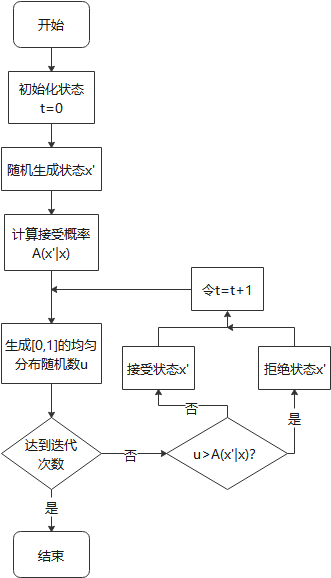


图 2‑1Metropolis算法流程图

## 二维Ising模型概述

Ising模型是统计物理学中研究最广泛的系统之一，它是在20世纪早期作为磁化和相关现象的模型发展起来的。

在统计物理中，定义Ising模型的物理概率分布是玻尔兹曼分布，其概率可表示为：

（2-1）

其中Z是配分函数，是用来描述一个统计系综中微观结构的分布状态的参数，是反温度函数，H是系统的哈密顿量。

Ising模型最初是一个离散的铁磁性模型[16]，由单个原子的磁矩或自旋及其相互作用能描述，其中每个晶格点自旋参数为，它们通过一个称为哈密顿量的函数对整体的能量行为做出贡献，整个系统的哈密顿量可写成：

（2-2）

其中<i, j>表示相邻的晶格点i、j，为晶格点的磁矩，哈密顿量的第一项表示晶格点之间自旋的交互作用，第二项表示外磁场对自旋相互作用的能量，本课题不考虑外磁场的能量作用，即第二项为0。

其中是一个交互作用的参数，其正负号决定了该Ising模型的磁性质：时，系统为铁磁性；时，系统为反铁磁性；时，晶格点i、j的自旋无交互作用。在本课题中，选择，系统具有铁磁性，便于计算。

## 预测参数的定义

本课题预测二维Ising模型的参数有能量、磁化强度、比热和磁化率四个参数，它们都与温度相关。

在本课题中，二维Ising模型忽略了外磁场的作用，且选取交互参数，且为一常数，系统为顺磁模型，故系统的能量为：

（2-3）

磁化强度是用来衡量物体磁性的物理量，该系统的磁化强度为：

（2-4）

磁化率为：

（2-5）

比热容是用来衡量系统的吸热或散热能力的，其定义为：

（2-6）

二维Ising模型是已知最早被证明存在相变的物理模型，其相变的物理意义是：系统处于低温的平衡态时，系统的自旋呈现有序状态；而当温度升高时，系统的自旋会出现一定的无序扰动；达到相变温度时，系统的自旋组态开始呈现无序状态[27]，温度越高，无序程度越高。因为温度会使其中一些晶格点自旋翻转，又影响周围晶格点的自旋，使整个系统的状态无序地发生改变，呈现出系统的无序性。本课题在预测二维Ising模型比热的同时，可以通过比热随温度变化的图像，找出相变点，预测相变的临界温度。并且对于Ising模型而言，系统的相变的临界温度是不变的，不与模型的初始状态相关[38]。

## 反向传播神经网络模型概述

反向传播（BP）算法是来自人工神经网络领域的多层前馈网络的监督学习方法[13]。前馈神经网络的灵感来自一个或多个称为神经元的神经细胞的信息处理。神经元通过其树突接受输入信号，这些树突将电信号向下传递到细胞体。轴突将信号传递到突触，突触是细胞轴突与其他细胞树突的连接。反向传播方法的原理是通过修改输入信号的内部权重以产生预期的输出信号来对给定函数建模。使用监督学习方法对系统进行培训，其中系统输出与已知预期输出之间的误差会显示给系统，并用于修改其内部状态。从技术上讲，反向传播算法是一种在多层前馈神经网络中训练权重的方法。因此，它需要定义一个或多个层的网络结构，其中一层完全连接到下一层。

BP网络的学习规律是：运用均方误差和梯度下降法对网络的连接权值进行更新修正[17]。该算法首先给网络的连接值取一个较小的值，然后选取一个训练样本来计算相对于该样本的误差梯度。

图 2‑2给出了一个简单的反向传播神经网络结构，包含一个输入层、一个输出层和一个隐藏层。

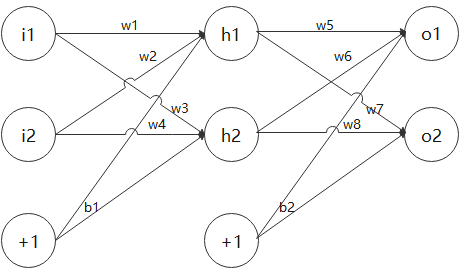


图 2‑2一个简单的BP神经网络结构

图 2‑1中i1、i2为神经元的输入，输入一般是对系统模型有较为关键影响的自变量；w1、w2、…、wj为连接各层神经元的权重；h1、h2为该结构中的隐藏层，o1、o2为神经元的输出层。由最便捷的线性加权求和可以求得各个神经元的净输入（以h为例）：

（2-7）

神经元h1的输出需经过隐藏层的激励函数得h1的输出，激励函数一般为sigmoid函数：

（2-8）

同理也可得到其他几个神经元的净输出和输出层的输出。

为了衡量实际输出与期望输出之间的差异，我们定义神经元的总误差能量为：

（2-9）

通过输入以及输出，计算误差，正向传播结束。接着计算反向传播，定义误差信号为网络的实际输出与期望输出之间的差值；在误差信号的反向传播中，信号从输出端逐层传播到输入层。迭代过程中，通过对权值和偏移值的连续修改，使网络的实际输出更接近期望值。

误差信号反向传播权值的更新规则[19]为（以为例）：

（2-10）

至此，一直更新至损失函数取得最小值或完成设置的迭代次数为止，在这个过程中，误差信号在不断减小。由此得到该神经网络中最佳的参数。

## 基于受限玻尔兹曼机的Ising模型实现

受限玻尔兹曼机是一种可通过输入数据集学习概率分布的随机生成神经网络[24]。训练该网络的方法是调整其模型参数，使得生成统计上与训练数据集相似的数据集的概率最大化，这也是通过随机梯度下降来实现的，与bp神经网络类似，对权重进行更新。

Alan和Roger（2018）[2]曾基于受限波尔兹曼机对二维Ising模型进行研究，使用蒙克卡罗技术生成数据集训练神经网络生成模型，由不同结构的浅层和深层模型生成自旋样品的能量测量。将测量值与训练模型所依据的数据的蒙特卡罗值进行比较。得出的其中一个结论为：在一定数量的网络资源下，Ising系统的最有效的表示是由一个只有一个隐藏层的浅层RBM给出的。

## 性能比较与本章小结

如此一来我们便从两个角度考量了Ising模型用于机器学习的可行性，可以看到bp网络相较于传统的受限玻尔兹曼机在学习速度和预测精度上有一定的优势。

本章具体阐述了模型所需的算法和机器学习方法的理论背景与计算模拟过程，详细说明了程序运算的流程，并画出流程图，同时具体说明了反向传播神经网络的理论方法及具体实现。最后介绍了基于受限玻尔兹曼机的Ising模型的构建，进一步说明了网络结构选取的原因与其有效性。

# 模型的设计

上一章论述了Ising模型以及所用算法和神经网络结构的理论背景与具体方程。本章详细给出了二维Ising模型的Metropolis算法的初始架构，接着对模型使用的反向传播算法的设计进行了具体的说明，详细介绍使用技术的技术原理和实现。

## 基于Metropolis算法的二维Ising模型设计

用Metropolis数值计算方法模拟Ising模型的演算过程[12]，图 2‑2给出了计算二维Ising模型的Metropolis算法流程图：

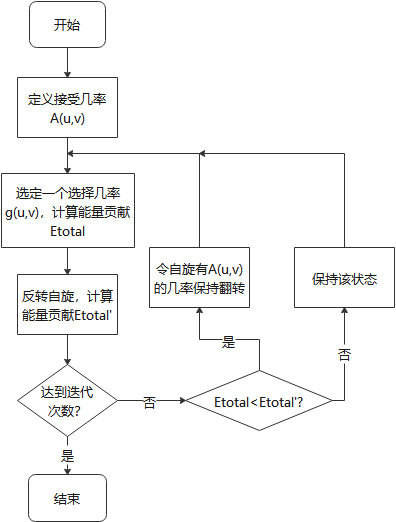


图 3‑1二维Ising模型的Metropolis算法流程图

其具体过程如下：

1. 用选择几率选择一个自旋初态，并计算所有与其相邻自旋作用的能量贡献；
2. 反转其自旋，再次计算所有与其相邻自旋作用的能量贡献；
3. 若该自旋的能量贡献下降，即，则保持这个状态；
4. 若该自旋的能量贡献下降，即，则令自旋有的几率保持翻转；
5. 重复步骤1，直至达到迭代次数。

## 反向传播算法实现设计

本课题设计的bp神经网络拥有一个输入层、一个输出层和一个隐藏层，其中输入层一维矩阵表示二维Ising模型的初始状态，输出层为四个需要预测的参数：能量、磁化强度、比热和磁化率。

从输入层到隐藏层的矩阵表示为：

（3-1）

经过隐藏层的激活函数见式（2-7）后，信号从隐藏层到输出层，该过程的矩阵表达式为：

（3-2）

经过输出层的激活函数后，计算该输出与期望值之间的误差，其中期望值由数值计算出的四个参数值给出。通过该误差，反向传播该信号，更新权值v、w，见式（2-10）的权值更新方程，不断进行上述操作进行迭代，使该误差越来越小，即可最终得到该网络模型的更新权值，达到训练的目的。

## 基于反向传播神经网络的Ising模型预测设计

反向传播算法是一种人工神经网络，是监督式的学习方法，可以比较有效地对一个物理参数进行预测。已知一个物理模型的初始状态及其物理过程后，得出这个物理模型的参数有多种方式，可以选择使用数值计算方法，如使用蒙特卡洛方法或积累量方法，同时，也可以与机器学习进行结合，使用受限玻尔兹曼机或神经网络。本课题选取反向传播神经网络的原因在于，BP算法的结构较为简单，只由三层网络组成，便于设计与实现；其次，使用其他复杂的神经网络可能会发生过拟合的可能；最后，对于Ising模型的数据预测选用由数据集训练的BP神经网络更加合适。

用神经网络架构对模型处于某一状态时的能量与比热进行预测，最关键的在于确定神经网络的输入层和输出层。将数值计算出的数据作为神经网络的数据集，即：用随机函数生成Ising模型的初始自旋状态，将其作为输入层，用蒙特卡洛方法计算出二维Ising模型的能量、比热、磁化率等参数，将其作为输出层。

把这些已知参数代入，经过n次迭代，不断地将权值和偏移值进行更新，使误差越来越小，使网络的输出值无限接近于期望输出值，由此得到该模型最佳的参数，使用这些参数对其他初始状态的Ising模型进行能量、比热的预测，并画出能量或比热随温度的变化曲线，与期望曲线相比较。

图 3‑2给出了上述描述过程的流程图：

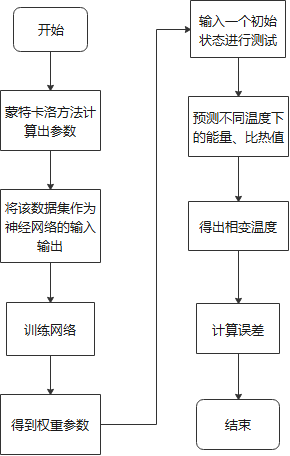


图 3‑2基于bp网络的Ising模型过程流程图

## 本章小结

本章详细说明了用矩阵的方式细致地描述了反向传播神经网络的设计过程，简单阐述了选择BP神经网络作为二维Ising模型方法的原因，并说明了基于Metropolis算法的二维Ising模型和基于反向传播神经网络的二维Ising模型预测实现设计，说明其中的联系与实现方法，并画出流程图。

# 模型实现与测试

上一章具体给出了模型的设计。本章详细地给出上述理论的实现过程与网络架构，介绍了使用神经网络对二维Ising模型比热预测的具体实现。首先对项目开发环境进行说明。接着分为三个小节，从模型的网络架构、训练框架的编写、测试框架的编写三个角度，对模型的实现与测试完整地进行了说明，并画出测试的数据图，与期望图对比。最后进行性能分析与比较。

## 开发环境

本项目的网络搭建使用NumPy。NumPy是Python中科学计算与数学的基础包，它提供多维数组对象以及数组的快速操作，可以很好地实现统计学的计算，以及搭建简单的神经网络。

模型的代码实现在个人计算机上完成。实验平台的硬件配置如表格 3‑1所示。

表格 3‑1系统实现硬件平台

|  |  |
| --- | --- |
| 名称 | 型号 |
| 中央处理器 | Intel(R) Core(TM)i5-7200U |
| 内存 | 4 GB |
| GPU显卡 | Intel(R) HD Graphics 620 |
| 操作系统 | Windows 10.0.17134.1304 |

开发工具为Jupyter Notebook，环境配置使用Python 3.6版本的Anaconda。实现过程中用到的部分Python包为：CUDA版本10.0.130，Cudnn版本7.6.5，Pytorch版本1.2.0，SciPy版本1.4.1，TensorboardX版本2.0，NumPy版本1.18.1，git版本2.23.0。

## 模型框架实现

本模型主要框架由两部分组成，一个是Ising模型的蒙特卡洛算法计算，另一个是反向传播神经网络的程序实现。

第一部分为二维Ising模型的Metropolis算法实现，定义的主要函数有：

*def init(L) return state*

*def E\_dimensionless(config,L) return (total\_energy/4)*

*def magnetization(config) return Mag*

*def MC\_step(config, beta) return config*

*def calcul\_energy\_mag\_C\_X(config, L, eqSteps, err\_runs) return (T,Energies,Magnetizations,SpecificHeats,Susceptibilities, delEnergies, delMagnetizations,M\_theoric, C\_theoric, delSpecificHeats, delSusceptibilities)*

第一部分的算法主要过程如下：首先对二维Ising模型的状态初始化，接着计算模型的初始总能量，然后用Metropolis算法更新状态，最后一个函数通过一定的迭代次数更新模型的状态，并计算出系统在该初始状态下的一些参数，如能量、磁化、比热、磁化率等。

其中的蒙特卡洛过程的函数为MC\_step(config, beta)，其程序中实现的具体过程为：对Ising模型的初始状态的二维矩阵进行每个元素的遍历，随机选取其中的一个晶格点，计算其能量，因为其周围上下左右四个晶格点的自旋作用，故对该晶格点的能量进行更新，直到对所有的晶格点进行更新完毕，该函数的具体代码实现为：

*def MC\_step(config, beta):*

*L = len(config)*

*for i in range(L):*

*for j in range(L):*

*a = np.random.randint(0, L) # looping over i & j therefore use a & b*

*b = np.random.randint(0, L)*

*sigma = config[a, b]*

*neighbors = config[(a+1)%L, b] + config[a, (b+1)%L] + config[(a-1)%L, b] + config[a, (b-1)%L]*

*del\_E = 2\*sigma\*neighbors*

*if del\_E < 0:*

*sigma \*= -1*

*elif np.random.rand() < np.exp(-del\_E\*beta):*

*sigma \*= -1*

*config[a, b] = sigma*

*return config*

第二部分实现反向传播神经网络算法，定义的主要函数有：

*def sigmods(z) return a*

*def forwordmd(X,W,V,B1,B2) return pred\_y,H*

*def Bpaugorith(Y,pred\_y,H,V,aph,W) return Vupdate,Wupdate,Errorterm*

第一个函数是激活函数sigmods函数，便于之后的矩阵运算；第二个是正向传播过程的函数，返回值为预测值；第三个函数是反向传播过程，用于计算误差，更新权重。

其中第二个正向传播过程的函数是基于式（3-1）、（3-2）的矩阵运算设计实现的，该函数的具体代码实现为：

*def forwordmd(X,W,V,B1,B2):*

*net1 = W.T\*X+B1*

*H = matrix(sigmods(np.array(net1))).T # 隐藏层单元*

*net2 = V.T\*H+B2*

*pred\_y = matrix(sigmods(np.array(net2))).T # 预测值*

*return pred\_y,H*

第三个反向传播函数式基于式（2-9）的误差公式和式（2-10）的权值更新公式进行设计实现的，该函数的具体代码实现为：

*def Bpaugorith(Y,pred\_y,H,V,aph,W):*

*Errorterm = 0.5\*(Y-pred\_y).T\*(Y-pred\_y)# 给出误差公式*

*# 计算输出单元的误差项*

*a1 = multiply(pred\_y-Y,pred\_y) # 矩阵对应元素相乘*

*a2 = multiply(a1,1-pred\_y)*

*Verror = H\*a2.T*

*# 计算隐藏单元的误差项*

*Werror = X\*(multiply(multiply(H,1-H),(V\*a2))).T*

*# 更新权重*

*Vupdate = V - aph\*Verror*

*Wupdate = W - aph\*Werror*

*return Vupdate,Wupdate,Errorterm*

主程序部分首先给出了两部分中的具体参数，如迭代次数、模型规模、学习率等等。接着随机生成模型的一个初始参数，调用第一部分的函数生成在这一种初始状态下的能量、比热等参数矩阵值。将这些值作为第二部分的输入，到一定的迭代次数后即可得到该bp神经网络对于二维Ising模型的最佳权值。

图 4‑1给出了模型实现的主函数流程图：

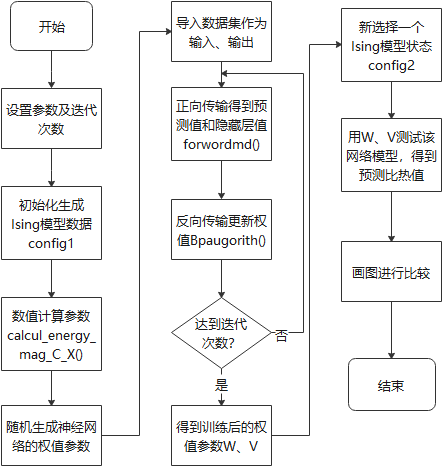


图 4‑1主函数实现流程图

## 模型训练实现

模型训练的主要过程如下：

1. 建立反向传播神经网络模型，并初始化二维Ising系统的状态为config，其中config的矩阵元素选取必须为1或-1。
2. 通过Metropolis算法的calcul\_energy\_mag()函数，计算得到参数的数据集。
3. 导入上一步的数据集，进行参数的读取，将config作为输入层矩阵，期望值参数作为输出层数据。
4. 运行神经网络结构程序，网络正向传播得到输出层。
5. 计算网络训练的误差参数。
6. 利用误差参数进行网络的反向传播，并进行权值的更新。
7. 训练过程中，实时输出预测参数的值，并在训练完成时保存最终的模型参数。

图 3‑1是该训练程序的实现流程图。

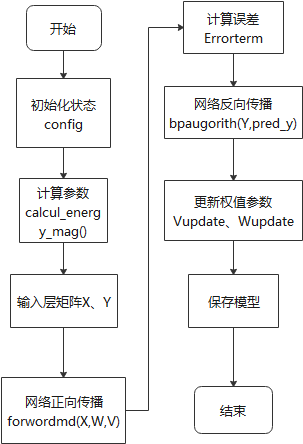


图 4‑2网络训练的主函数流程图

## 模型测试实现

模型的预测具体实现过程为：

1. 随机选择二维Ising模型的一个初始状态config2，其中config的矩阵元素选取必须为1或-1，且最好不与config1相同。
2. 使用训练网络结构后的权值参数作为神经网络的权值参数导入，进行该模型的测试。
3. 利用正向传播函数计算此时状态下模型的输出，即config2状态下二维Ising模型的能量、比热等参数。
4. 将这些预测参数保存，并和期望参数画在一张图上进行比较，计算该网络模型的误差。

模型的测试需要读入本模型一些随机的初始状态，预测该状态下的参数值以及相对于期望值的误差。图 3‑2展示了测试程序的总体架构主函数流程图。

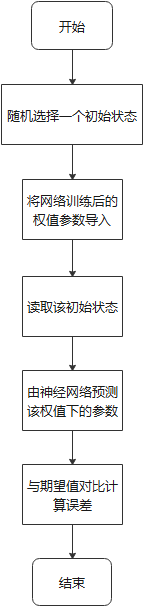


图 4‑3测试程序的主函数流程图

## 结果呈现与分析

首先生成二维Ising模型的能量、比热的数据集。通过Metropolis蒙特卡洛方法数值计算出Ising模型一定的二维初始状态下的能量、比热等参数值。

选择模型初始状态为：

（3.1）

其中初始矩阵元素的取值均为1或-1，分别表示该二维Ising模型的晶格点原子磁矩的自旋方向为向上或向下。

设置蒙特卡洛过程的迭代次数，并将初始状态config1导入calcul\_energy\_mag\_C\_X函数计算出二维Ising模型的能量、比热等参数值。图 3‑3给出了比热的误差棒图：

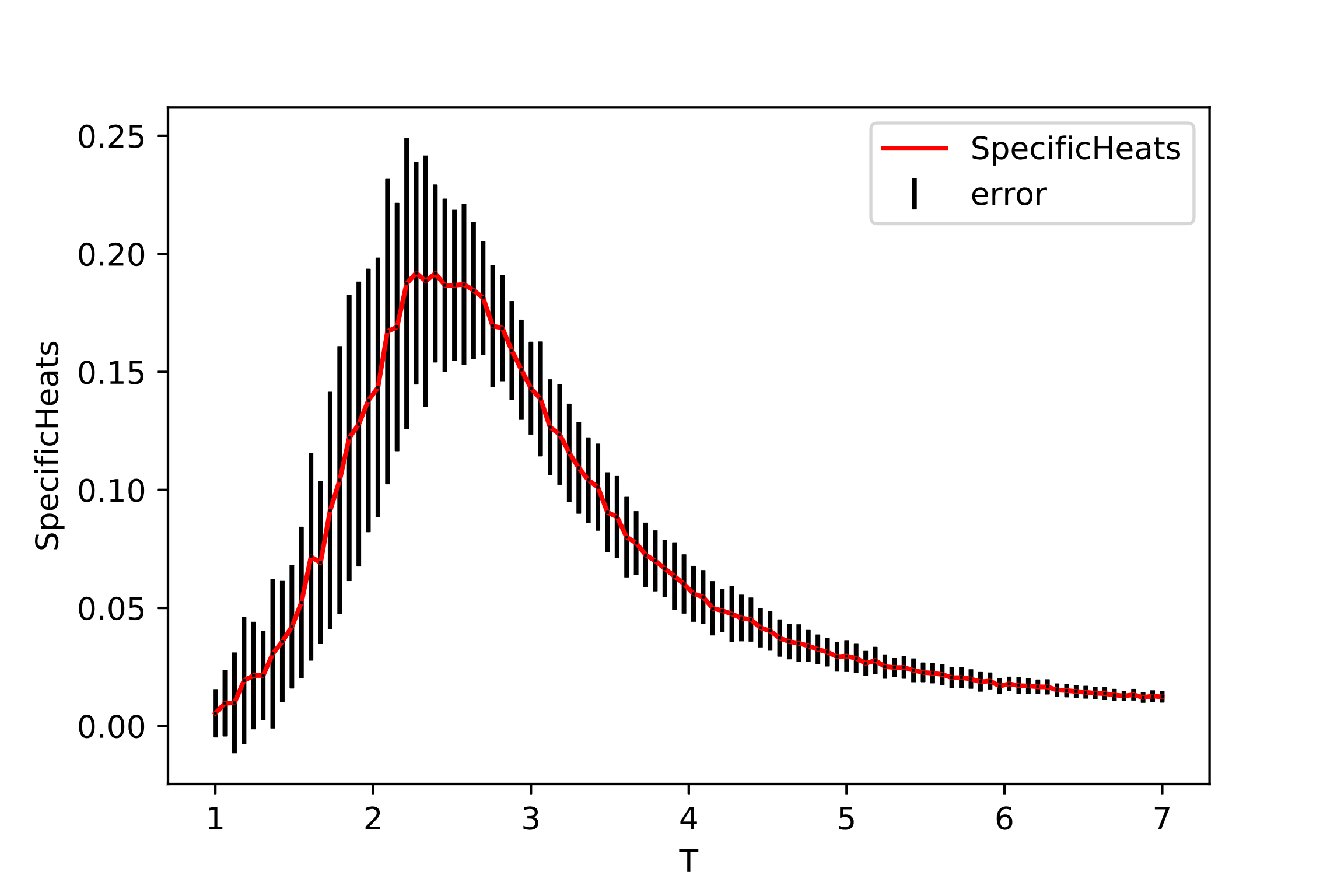


图 4‑4数值计算比热的误差棒

将数值计算得到的四个参数的数据集导入神经网络中进行训练，设置迭代次数，学习率，初始状态仍为config1。开始监督式的训练过程，达到迭代次数后，将训练完成后预测的参数值随温度的变化画出曲线图 3‑3训练时预测和期望的比热与温度的关系：

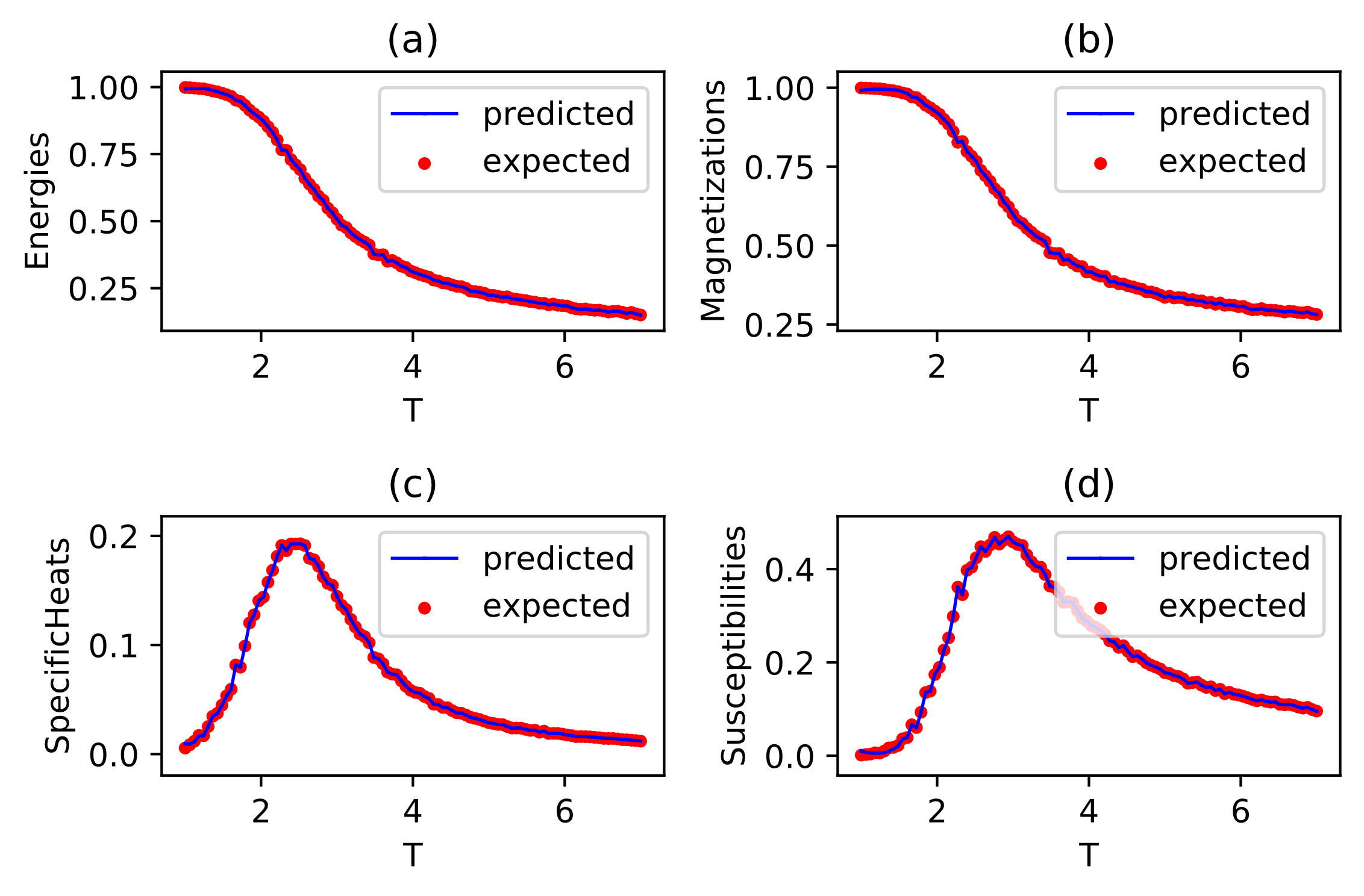


图 4‑5训练结束后预测参数值和期望值与温度的关系

图 3‑3给出了训练神经网络完成后预测的能量、磁化、比热、磁化率与期望参数值随温度的变化关系，其中的子图(a)、(b)、(c)、(d)分别表示参数为能量、磁化、比热、磁化率的曲线对比图。其中，能量和磁化都是随温度升高而降低的，这是符合物理规律的，因为随着温度升高，由于热骚动造成的碰撞会影响磁矩的排列整齐程度，能量会下降，顺磁性会降低；而随温度的升高，比热和磁化率是先升高后降低的，这是出现了相变现象，其中的极大值点即为该初始状态下二维Ising模型的相变点。

图 3‑4训练结束后的误差随温度的变化给出了训练结束后总误差随温度的变化关系。可以明显地看出，该模型的误差较小，数量级为，且在温度较大时误差基本为0。并且图中突变的误差极大值所对应的点为模型的相变点，这符合本课题所描述的物理模型。

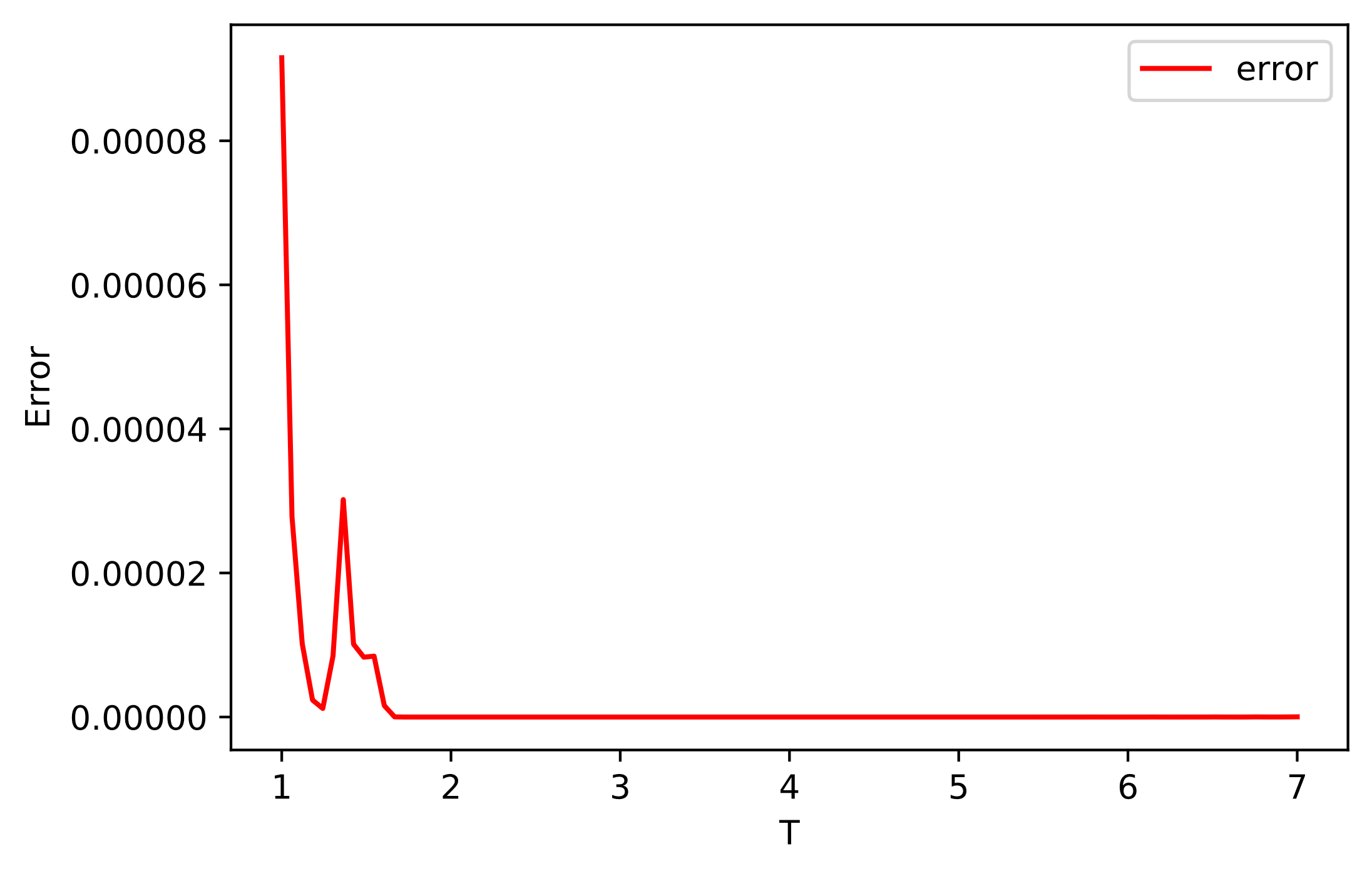


图 4‑6训练结束后的误差随温度的变化

下面进行测试阶段，选取任意的初始状态，这里选取的初始状态为：

（3.2）

使用训练结束后的更新权值作为该测试的权值，调用向前传播函数预测参数。

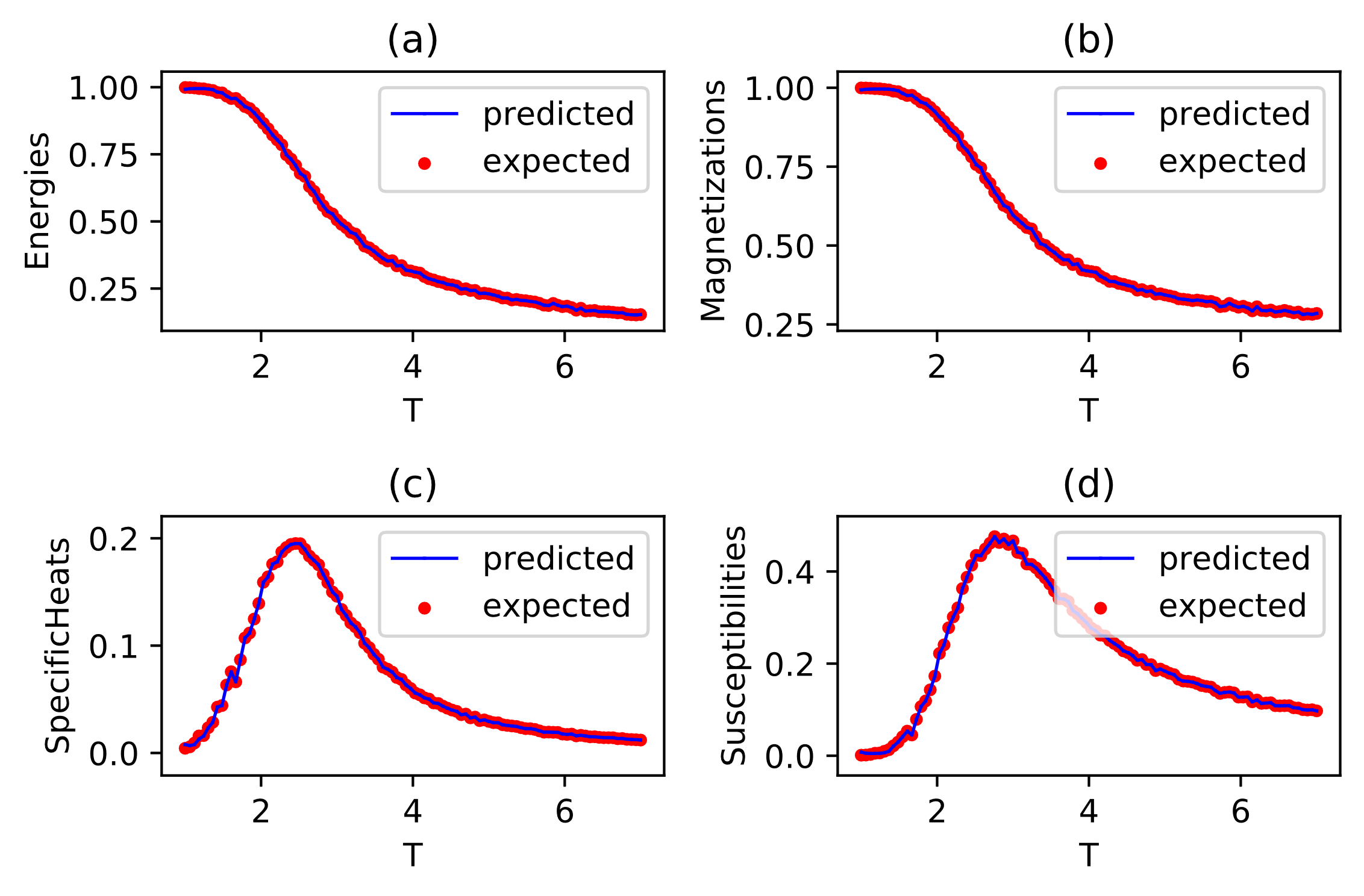


图 4‑7模型测试时预测和期望的参数值与温度的关系

图 3‑4给出了一定初始状态下该模型预测的参数值与期望参数值随温度变化的曲线对比图。其趋势基本与期望值的曲线趋势相同，且符合得较好。表格 3‑2给出了测试时预测的参数值与期望值的误差：

表格 3‑2 测试时预测参数误差

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 能量 | 磁化强度 | 比热 | 磁化率 |
| 误差 | 4.65840e-05 | 1.62299e-04 | 1.71803e-05 | 1.74058e-04 |

其中误差的数量级最大为，说明该模型的预测功能的可行性。

另外，通过图 3‑5中（c）图中的极大值点，可求得该初始状态下的二维Ising模型的相变温度约为。

综合上述实验结果与分析，本课题最终实现的神经网络的模型可基本实现对二维Ising模型的参数预测与相变温度的预测，表明本项目实现的网络架构是有效的。

## 性能对比与分析

本课题预测二维Ising模型比热的误差为1.71803e-05，所属数量级较小，且图 3‑6模型测试时预测和期望的参数值与温度的关系中未出现预测偏差明显的点，证明了预测的正确性与有效性。

由表格 3‑2 测试时预测参数误差和图 3‑5训练结束后的误差随温度的变化可以看出，相较于使用Metropolis算法的蒙特卡洛数值计算方法，基于bp神经网络的模型对于二维Ising模型的比热预测误差更大，但差别不大，可以忽略。

在一维、二维和三维的Ising模型中，利用累积量方法可以从小晶格的磁化涨落中确定李阳零点[39]，从而预测相变的发生。该模型能正确地预测Ising模型的磁化强度，虽然存在偏差，但该模型适用于三维Ising模型和许多一阶相变系统。

基于混合蒙特卡洛（HMC）模拟的Ising模型[40]演示了如何用连续变量来模拟具有离散自由度的系统。Metropolis算法相比于该模型对于计算Ising模型更加精确，且该算法在效率上比Metropolis算法更低效，但在未来，HMC算法可以通过更有效的积分器和参数的不同选择来进行优化，加快Ising模型的计算。

综上所述，本课题的设计与结果比较成功，误差较小，模型具有可靠性。

## 本章小结

本章首先介绍了程序架构时使用的神经网络框架，以及项目的开发环境。随后详述了基于bp神经网络的Ising模型的具体实现，然后给出了训练与测试程序时的执行流程图。接着给出了完整的模型测试方案，对二维Ising模型进行训练与测试，得到由预测参数值和期望参数值组成的曲线比较图，并计算其误差，证明了模型的可行性。最后与二维Ising模型的其他方法进行性能对比与分析，发现本课题的误差较小，实现性较高。

# 总结与展望

本课题主要背景是二维Ising模型以及神经网络领域，分析了当前存在的几个技术难点，对于预测一些参数、高维度等问题，给出一定的解决方案。本课题做了如下几点研究工作：

1. 用数值计算的方法较好地估计了二维Ising模型不同温度下的比热，架构反向传播神经网络，并用其训练网络模型，以预测模型各种初始状态的比热、能量、磁化强度和磁化率与温度的关系。
2. 训练后的神经网络模型较好地预测了二维Ising模型不同初始状态下不同温度时的比热，以及系统相变的临界温度。

同时，本课题的模型架构中还存在一些问题尚未解决，但统计物理学与机器学习领域的相互交流与促进一直存在，对Ising模型的数据预测方法仍有许多待改进之处。

在未来的工作中，需要改进的地方总结如下：

1. 增加神经网络中的隐藏层，探究隐藏层的多少是否对模型的性能产生影响，深层架构的网络模型是否对模型参数的预测更有说服力。
2. 可以架构卷积神经网络，对二维系统的情形描述得更加正确与精准，提高网络的效率，减少误差。
3. 讨论三维的Ising模型，需要考虑更高精度的蒙特卡洛方法[34][35][37]，以及机器学习方法[36]。

参考文献

1. Talbi M , Ftima S B , Cherif A . [IEEE 2017 International Conference on Control, Automation and Diagnosis (ICCAD) - Hammamet, Tunisia (2017.1.19-2017.1.21)] 2017 International Conference on Control, Automation and Diagnosis (ICCAD) - Speech modulation for image watermarking[J]. 2017:522-527.
2. Morningstar A , Melko R G . Deep Learning the Ising Model Near Criticality[J]. Journal of Machine Learning Research, 2017.
3. Mehta P , Bukov M , Wang C H , et al. A high-bias, low-variance introduction to Machine Learning for physicists[J]. Physics Reports, 2018, 810.
4. Mcculloch W S , Pitts W H . A logical calculus of the ideas immanent in nervous activity[J]. The Bulletin of Mathematical Biophysics, 1988, 5:115-133.
5. D. E. Rumelhart, G. E. Hinton, R. J. Williams. Learning Internal Representations by Error Propagation[M]// Neurocomputing: foundations of research. MIT Press, 1988.
6. Ethem Alpaydin. Introduction to Machine Learning (Adaptive Computation and Machine Learning)[M]// Introduction to machine learning /. MIT Press, 2004.
7. Hinton, G. E, Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks.[J]. Science, 2006.
8. Baldi P , Sadowski P , Whiteson D . Searching for Exotic Particles in High-Energy Physics with Deep Learning[J]. Nature Communications, 2014, 5:4308.
9. Rousseau D , Adam-Bourdarios C , Cowan G , et al. The Higgs Machine Learning Challenge[J]. Higgs Machine Learning Challenge Visits Cern, 2015, 664(7).
10. Beichl I , Sullivan F . The Metropolis Algorithm[J]. Computing in ence & engineering, 2000, 2(1):P.65-69.
11. Metropolis N, Rosenbluth A W, Rosenbluth M N, et al. Equation of state calculations by fast computing machines[J]. The journal of chemical physics, 1953, 21(6): 1087-1092.
12. 杨源杰, 黄道. 人工神经网络算法研究及应用[J]. 华东理工大学学报(自然科学版), 2002(05):115-118.
13. Sonsin A F , Cortes M R , Nunes D R , et al. Computational Analysis of 3D Ising Model Using Metropolis Algorithms[J]. Journal of Physics Conference, 2015, 630:012057.
14. SAS Software Ltd, Marlow, Derek Powell. Neural Networks and Statistical Models[J]. 1994.
15. Castañeda-Marroquín C, de la Puente A O, Alfonseca M, et al. A threaded Java concurrent implementation of the Monte-Carlo Metropolis Ising model[C]//International Work-Conference on the Interplay between Natural and Artificial Computation. NIH Public Access, 2009, 2009: 103.
16. 刘天舒. BP神经网络的改进研究及应用[D]. 东北农业大学.
17. Flack A, Bertini B, Prosen T. Statistics of the Spectral Form Factor in the Self-Dual Kicked Ising Model[J]. arXiv preprint arXiv:2009.03199, 2020.
18. Yi J, Wang Q, Zhao D, et al. BP neural network prediction-based variable-period sampling approach for networked control systems[J]. Applied Mathematics and Computation, 2007, 185(2): 976-988.
19. Varsamopoulos S, Bertels K, Almudever C G. Designing neural network based decoders for surface codes[J]. arXiv preprint arXiv:1811.12456, 2018.
20. Efthymiou S, Beach M J S, Melko R G. Super-resolving the Ising model with convolutional neural networks[J]. Physical Review B, 2019, 99(7): 075113.
21. Li J, Cheng J, Shi J, et al. Brief introduction of back propagation (BP) neural network algorithm and its improvement[M]//Advances in computer science and information engineering. Springer, Berlin, Heidelberg, 2012: 553-558.
22. 蒋鹏飞. 温度对不同初始状态 ising 模型磁化强度和磁化率的影响[J]. 河南科技, 2014 (9X): 278-279.
23. Hinton G E. Deep belief networks[J]. Scholarpedia, 2009, 4(5): 5947.
24. Ising E. Contribution to the theory of ferromagnetism[J]. Z. Phys, 1925, 31(1): 253-258.
25. Onsager L. Crystal statistics. I. A two-dimensional model with an order-disorder transition[J]. Physical Review, 1944, 65(3-4): 117.
26. 刘策军, 郑有因. 二维伊辛模型自旋状态图样的蒙特卡罗模拟[J]. 华南农业大学学报, 1995, 016(004):101-105.
27. 张志东. 伊辛模型的研究进展简介[J]. 自然杂志, 2008(02):94-98.
28. Bragg W L, Williams E J. The effect of thermal agitation on atomic arrangement in alloys[J]. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1934, 145(855): 699-730.
29. Shockley W. Theory of order for the copper gold alloy system[J]. The Journal of Chemical Physics, 1938, 6(3): 130-144.
30. Montroll E W. Statistical mechanics of nearest neighbor systems[J]. The Journal of Chemical Physics, 1941, 9(9): 706-721.
31. Kaufman B. Crystal statistics. II. Partition function evaluated by spinor analysis[J]. Physical Review, 1949, 76(8): 1232.
32. Wilson K G. Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture[J]. Physical review B, 1971, 4(9): 3174.
33. Tsypin M M, Blöte H W J. Probability distribution of the order parameter for the three-dimensional Ising-model universality class: A high-precision Monte Carlo study[J]. Physical Review E, 2000, 62(1): 73.
34. Janke W, Villanova R. Ising model on three-dimensional random lattices: A Monte Carlo study[J]. Physical Review B, 2002, 66(13): 134208.
35. Zhang R, Wei B, Zhang D, et al. Few-shot machine learning in the three-dimensional Ising model[J]. Physical Review B, 2019, 99(9): 094427.
36. Xu J, Ferrenberg A M, Landau D P. High-resolution Monte Carlo study of the order-parameter distribution of the three-dimensional Ising model[J]. Physical Review E, 2020, 101(2): 023315.
37. 蒋鹏飞. 温度对不同初始状态 ising 模型磁化强度和磁化率的影响[J]. 河南科技, 2014 (9X): 278-279.
38. Deger A, Brange F, Flindt C. Lee-Yang theory, high cumulants, and large-deviation statistics of the magnetization in the Ising model[J]. arXiv preprint arXiv:2006.15125, 2020.
39. Ostmeyer J, Berkowitz E, Luu T, et al. The Ising Model with Hybrid Monte Carlo[J]. arXiv preprint arXiv:1912.03278, 2019.
40. Rohit Kumar Sachan, Ayush Nigam, Avinash Singh, et al. Optimizing Basic COCOMO Model Using Simplified Genetic Algorithm[J]. Procedia Computer Science, 2016, 492-498.