

Applicability and Implementation of Self-Supervised Ensemble Learning in Phase Transition Classification

李彧(S10210036)

指導教師:吳桂光

*Email: s10210036@thu.edu.tw

摘要

本研究的主要目的是利用Self-supervised ensemble learning (SSEL)的方法來探索clock model中的相變。我嘗試利用全連接神經網路 (Fully Connected Neural Network) 作為模型架構進行訓練。並透過使用SSEL方法通過分析模型輸出的波動來準確地檢測相變點，之後希望透過此方法，調整模型中的參數，來探索其在potts model、Ising model或是更廣泛的相變識別和分類問題中的適用性。

研究計畫內容

(一) 研究動機與研究問題

SSEL模型與文獻[1]的自我監督學習 (SSL) 模型有些許的不同，雖然SSL也可以用來識別相變，但無法像SSEL模型那樣分類其屬性或類型。這兩種自我監督方法的根本區別在於它們的輸入特徵，在SSEL中，它來自原位測量，而在SSL中，它是在集合平均後獲得的某種物理量。前者 (原位測量) 確實包含了關於物質狀態的更多基本信息，使得SSEL能夠通過模型輸出的波動，即使在訓練階段 (參見Figure 1.) 也能識別相變。

會選擇此方法來探索clock model中的相變，是因為這種自監督集成學習方法不依賴於對底層理論模型的特定假設，並且僅通過實驗測量就能提取有價值的物理系統信息。此外，該方法還包含了比以前的機器學習方法更豐富的信息，並且可以應用於高維量子系統和其他複雜的物理現象。此方法可能有助於擴展對相變分類的理解，並為未來的研究提供新的方法和思路。

(二) 文獻回顧與探討

SSEL方法文獻[2]對物理學的重要性在於它提供了一種新的方法來識別和分類不同類型的相變，包括一階、二階和BKT相變。這對於理解物質的性質和行為以及開發新的物理模型和理論具有重要意義。此外，這種方法還有望應用於更高維度的量子系統和其他多體物理現象的研究。目前的研究進展表明，這種自監督集成學習方法能夠準確識別和分類不同類型的相變，並且在經典和量子系統中都表現出良好的應用前景。

重點是SSEL方法通過神經網路輸出的波動特性來識別相變點，並且能夠在不同的訓練階段保持穩定的性能。這表明了SSEL方法在研究相變性質方面具有重要的應用價值。

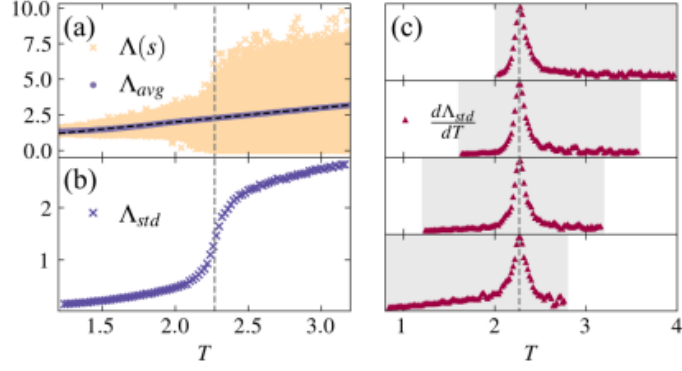


Figure 1. 透過SSEL方法輸出的波動特性來識別相變點。圖片來源:文獻[2]。

(三) 研究方法及步驟

在文獻[2]自監督集成學習 (SSEL) 方法中, 透過公式(1)對於給定的自旋構型集合 s , 通過神經網絡輸出的平均值與溫度之間的關係。

$$\Lambda_{avg}(S, T) = \frac{1}{N} \sum_{s \in S} \Lambda(s) \quad (1)$$

其中, $\Lambda(s)$ 示神經網絡對於自旋構型 s 的輸出, $E(s)$ 是給定自旋構型 s 的系統總能量, T 表示溫度, Z 是分配函數。這個公式用於計算神經網絡輸出和溫度之間的關係。

再使用公式(2)評估神經網絡輸出和溫度之間誤差的損失函數。其中, $\Lambda_{avg}(\tilde{S}_M, T_m)$ 表示在給定溫度 T_m 下, 神經網絡對於自旋構型集合 \tilde{S}_M 的平均輸出, T_m 表示溫度。這個損失函數用於評估模型性能作用。

$$L(\tilde{s}_1, \dots, \tilde{s}_M) = \frac{1}{M} \sum_{m=1}^M |\Lambda_{avg}(\tilde{S}_M, T_m) - T_m|^2 \quad (2)$$

最後使用公式(3)用於計算標準差, 以評估神經網絡輸出的波動程度。其中, $\Lambda_{std}(T)$ 代表標準差, N 代表樣本數量, s 代表樣本, $\Lambda(s)$ 代表神經網絡對於樣本的輸出, $\Lambda_{avg}(S, T)$ 代表在溫度 T 下樣本集合 S 的平均輸出。

$$\Lambda_{std}(T) \equiv \left[\frac{1}{N} \sum_{s \in S} |\Lambda(s) - \Lambda_{avg}(S, T)|^2 \right]^{1/2} \quad (3)$$

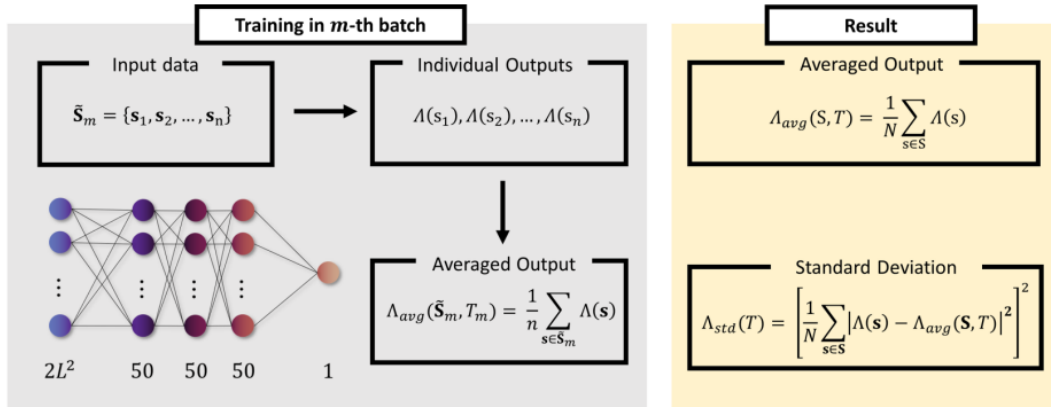


Figure 2. 文獻[2]中所提出的SSEL神經網路結構示意圖。圖片來源:文獻[2]。

1. q=2 clock model設置

我這裡選用q=2 clock model spin數據資料, 資料內容是80000筆的溫度, 然後每一個溫度都有對應自旋, 自旋為 $s_i \in \{+1 = \uparrow, 0 = \downarrow\}$, $i = 1, 2, \dots, N$, 模型尺寸單邊晶格數為 $L=20$, 總晶格數為 $N = L^2 = 400$ 。總共有80個不同的溫度, 每個溫

度都有相同的1000個溫度，並且我溫度範圍只取 $T=[1.05,1.10,\dots,3.8]$ 。再來我開始刪除不需要的列和行，將資料重新塑形為合適的格式、並將資料分組。這些步驟的目的是為了將原始資料轉換成可供模型訓練的格式。並把對應溫度下的自旋組態可視化如Figure 3。

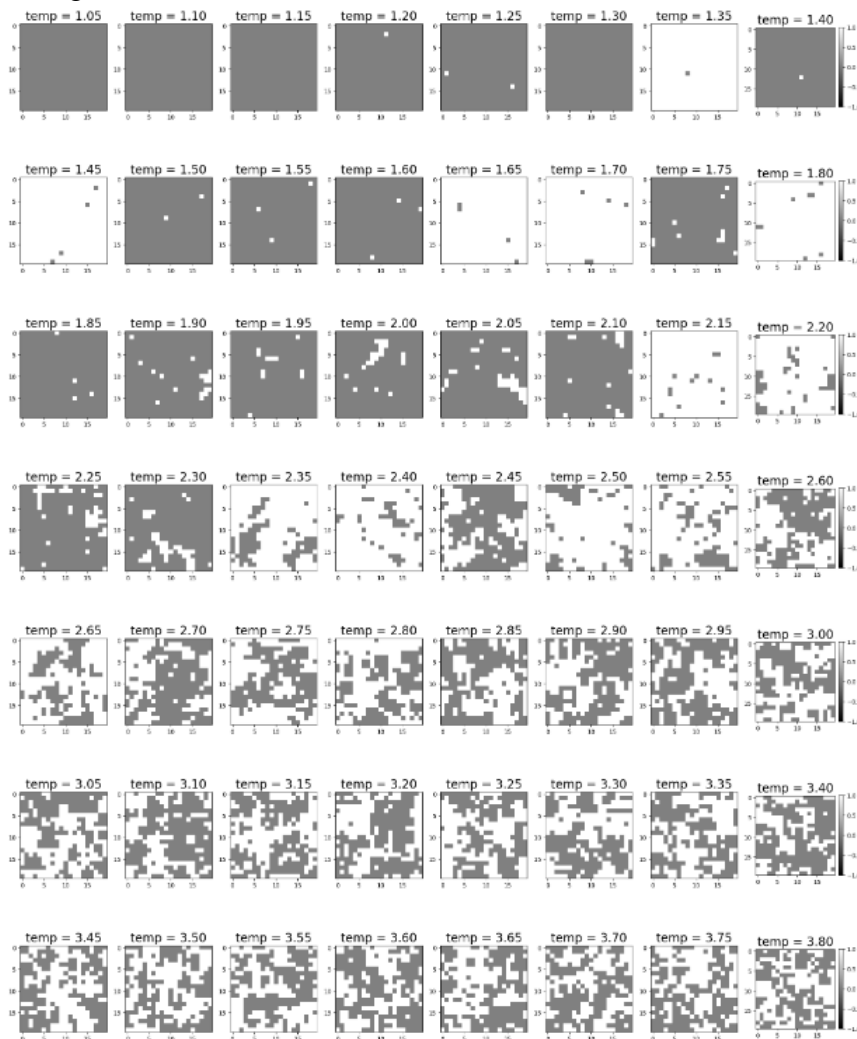


Figure 3. 把對應溫度下的自旋組態可視化。

2. Fully Connected Neural Network架構及運作方法

我使用的是Fully Connected Neural Network，它的架構是由三個隱藏層和一個輸出層組成。每一個隱藏層都使用SeLU文獻[3]作為激活函數，而輸出層則使用ReLU作為激活函數。並且使用迴圈的方式製作50個相同的子模型。製作出50個相同的子模型後，我使用concatenate()函數將它們連接起來，這種方法常用於模型集成（Model Ensemble），將多個模型的學習結果結合起來，以提高預測的準確性或穩健性。最後使用Lambda()函數創建一個層，該層的功能是對上一步連接後的50個子模型輸出進行平均。這一步的用意是為了組合多個模型的輸出，提高模型的泛化能力。

(四) (現有)結果與討論

透過損失函數對模型評估，用程式語法進行數據可視化，並且把重新塑形好的資料帶入架設好的模型裡面進行訓練，透過損失函數作圖Figure 4.，可以發現

跟Figure 5.文獻[2]的損失函數圖相比，我模型的損失函數收斂的地方相對太高，這也可能是導致我後面繪製原始數據點跟文獻[2]有異的主要原因。

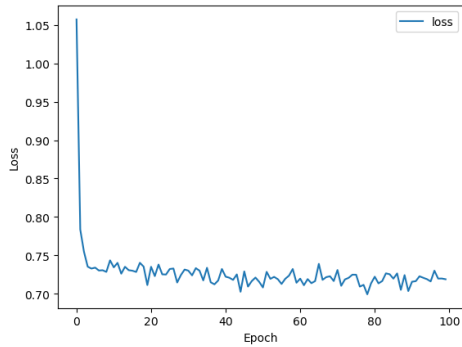


Figure 4. 訓練過程中誤差程度。

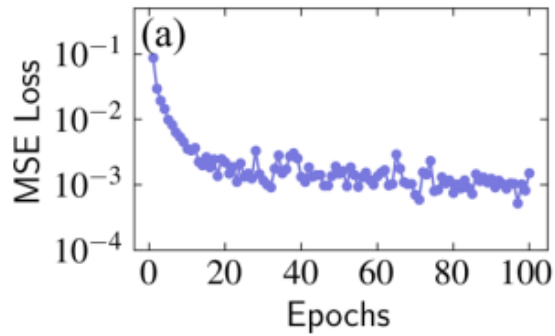
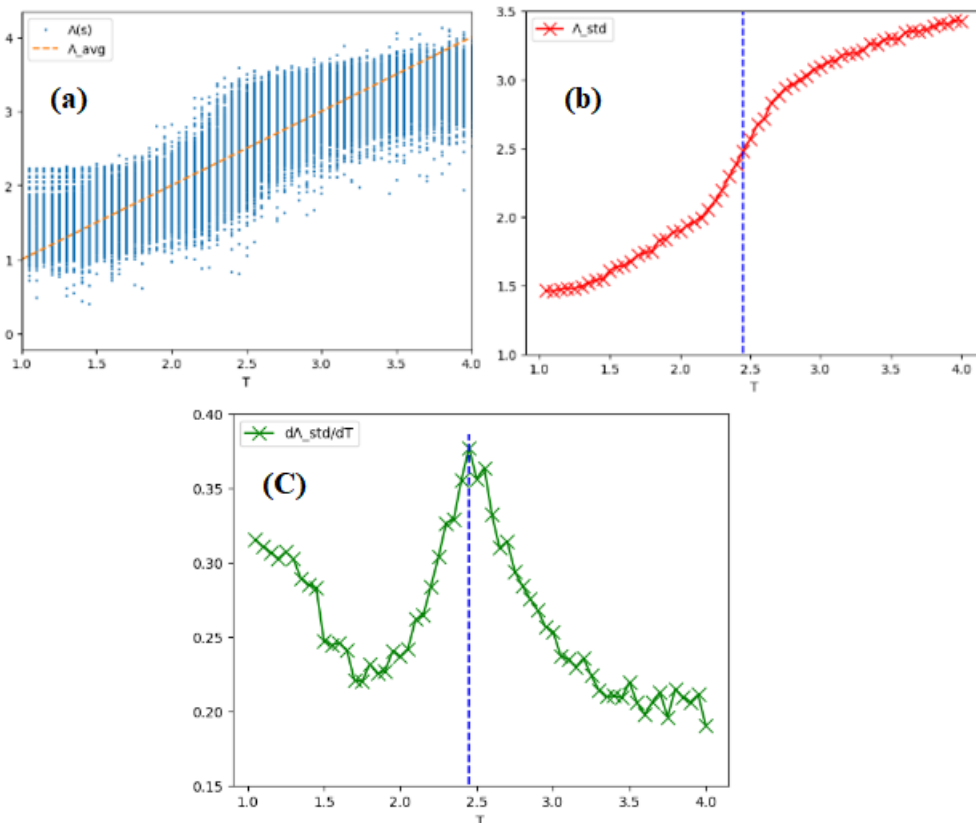


Figure 5. 文獻[2]訓練過程中誤差程度。

以下是 $L=20$ 的 $q=2$ clock model的SSEL結果。垂直虛線表示臨界溫度 $T_c \approx 2.4$ 。(a)來自模型的個別輸出 $\Lambda(s)$ (藍點)，對於給定的溫度和自旋配置。由於我們損失函數的約束，它們的平均值 Λ_{avg} 必須等於溫度 (橘色虛線)；請參見方程(2)。(b)計算模型輸出的標準差 Λ_{std} (紅色叉)，其在方程(3)中定義。(c) Λ_{std} 的第一溫度導數在臨界溫度 T_c 附近有一個顯著的峰值。



(五) 預期結果

1. 儘管損失函數要求平均輸出 Λ_{avg} 等於溫度 (參見方程(2))，但個別輸出可以是任何值而無需嚴格約束。實際上，隨著溫度的升高，其分佈變得更寬，反映出在較高溫度下的Boltzmann weight變得更加廣泛，但我目前個別的輸出沒有達到我預期的結果，所以接下來我會調整我的模型和分組好的數據，讓我的輸出可以達到我的預期。

2. 由文獻[4]可以知道，標準差對溫度的最大變化率位置很好地與已知的臨界溫度 $T_c \approx 2.269$ (垂直虛線)良好對齊，因此可作為相變的可靠指標。但我模型跑出來的臨界溫度(圖(c)垂直虛線表示臨界溫度)是 $T_c \approx 2.4$ ，所以我會再重新調整我模型架構和資料預處理的部分，讓我的數據能成功達到我的預期結果，之後也會把potts model或是Ising model的數據帶入我的模型做更多的嘗試。

(六) 參考文獻

- [1] C.-T. Ho and D.-W. Wang, Robust identification of topological phase transition by self-supervised machine learning approach, [New J. Phys. 23, 083021 \(2021\)](#).
- [2] Chi-Ting Ho and Daw-Wei Wang, Self-supervised ensemble learning: A universal method for phase transition classification of many-body systems, [Phys. Rev. Research 5, 043090 \(2023\)](#)
- [3] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, [Self-normalizing neural networks](#), Advances in Neural Information Processing Systems, edited by I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., 2017), Vol. 30
- [4] R. J. Baxter, Exactly Solved Models in Statistical Mechanics (Elsevier, Amsterdam, 2016)