Declaration: I declare that I have personally prepared this assignment. The work is my own, carried out personally by me unless otherwise stated and has not been generated using paid for assessment writing services or Artificial Intelligence tools unless specified as a clearly stated approved component of the assessment brief.All sources of information, including quotations, are acknowledged by means of the appropriate citations and references.I declare that this work has not gained credit previously for another module at this or another University, save for permitted elements which formed part of an associated proposal linked directly to this submission.

I understand that plagiarism, collusion, copying another student and commissioning (which for the avoidance of doubt includes the use of essay mills and other paid for assessment writing services, as well as unattributed use of work generated by Artificial Intelligence tools) are regarded as offences against the University's Assessment Regulations and may result in formal disciplinary proceedings. I understand that by submitting this assessment, I declare myself fit to be able to undertake the assessment and accept the outcome of the assessment as valid.

Student signature: Kevin Duah-Anim

Date: 30th September 2024

Unsupervised Machine Learning for Detecting and Characterising Phase Transitions in the 2D Ising Model

Kevin Duah-Anim

230313356

Contents

1 Introduction

1.1 Background

Phase transitions are fundamental processes in physics, representing a system's transformation from one state of matter to another due to variations in external conditions such as temperature, pressure, or electromagnetic fields (Goldenfeld, 2018). These transformations are classified into two main categories: first-order and second-order phase transitions. First-order transitions are characterised by abrupt changes in thermodynamic properties—such as energy, volume, or entropy—and involve the absorption or release of latent heat. For example, during the boiling of water, a system undergoes a first-order transition from liquid to gas while maintaining a constant temperature (Stanley, 1971). In contrast, second-order phase transitions exhibit continuous, gradual changes in the system's properties, with no latent heat involved. At the critical point of a second-order transition, physical quantities such as specific heat, magnetic susceptibility, and correlation length diverge, signalling long-range correlations and the emergence of critical phenomena (Kardar, 2007).

Understanding phase transitions is a cornerstone of both theoretical and applied physics. Theoretically, they offer profound insights into symmetry breaking, critical behaviour, and universality, which are foundational concepts in statistical mechanics and condensed matter physics (Stanley, 1971). Prac-

tically, phase transitions are critical for advancements in fields ranging from material science to quantum computing, where the control and manipulation of phases play a pivotal role in technological innovation (Sethna, 2006). Consequently, the study of phase transitions not only deepens our comprehension of natural phenomena but also drives significant progress in diverse scientific and engineering domains.

1.2 First-Order Phase Transitions

First-order phase transitions are distinguished by discontinuities in thermodynamic quantities at the transition point. During these transitions, a system absorbs or releases latent heat, which manifests without a corresponding change in temperature. Such transitions also involve a sudden change in volume, entropy, or other thermodynamic properties (Landau & Lifshitz, 1980). For example, the melting of ice involves a sharp change in volume and heat capacity as the system transitions from solid to liquid. During this process, energy is absorbed, driving the transition while maintaining a constant temperature (Huang, 1987).

Common examples of first-order phase transitions include melting, boiling, and sublimation. Each process involves a significant rearrangement of the system's microscopic structure, driven by the absorption or release of latent heat (Atkins, 2010). These transitions often exhibit phase coexistence, where different phases can exist simultaneously at equilibrium, such as the co-occurrence of liquid and gas phases during boiling. Metastability also plays a role in first-order transitions, where systems can remain in a non-equilibrium state for extended periods before transitioning to a more stable

phase (Gibbs, 1961). Another critical feature of first-order transitions is hysteresis, a phenomenon where the system's state depends on its history, as seen in the magnetisation behaviour of ferromagnetic materials when an external magnetic field is applied and cycled (Kittel, 2005).

1.3 Second-Order (Continuous) Phase Transitions

Second-order phase transitions, or continuous phase transitions, differ from first-order transitions in that they occur without abrupt changes in thermodynamic properties or latent heat. Instead, second-order transitions involve the continuous variation of an order parameter, which characterises the degree of order in the system. Near the critical point, physical quantities such as specific heat and magnetic susceptibility diverge, reflecting long-range correlations and critical phenomena (Goldenfeld, 2018). A classic example of a second-order phase transition is the ferromagnetic-to-paramagnetic transition, where magnetisation gradually decreases to zero as the temperature rises above the Curie point (Stanley, 1971). Another notable example is the superfluid transition in helium-4, where the fluid exhibits zero viscosity below a critical temperature (Donnelly, 2009).

Second-order transitions occur at critical points, where the system's thermodynamic behaviour undergoes qualitative changes without discontinuities in the first derivatives of the free energy (Fisher, 1974). Near these points, physical properties often follow power-law behaviour, necessitating advanced frameworks such as renormalization group theory to explain the critical behaviour and scaling laws (Cardy, 1996). Such phenomena exemplify the complexity of second-order transitions, making them a central focus in modern statistical

mechanics.

1.4 Critical Phenomena

In the context of phase transitions, critical phenomena refer to the divergent behaviour of physical quantities near the critical point. These behaviours are described by critical exponents—such as α, β, γ, δ, ν, and η—which quantify how observables such as specific heat, magnetisation, and susceptibility diverge as the system approaches the critical point (Stanley, 1971; Pelcovits, 1984). For example, α describes the divergence of specific heat, while β quantifies the growth of the order parameter (Stanley, 1971).

The scaling laws that govern these exponents arise from the concept of scale invariance at the critical point. These laws—such as $\alpha + 2\beta + \gamma = 2$ and $\nu = \frac{1}{d}$ (where d is the spatial dimension)—provide a framework for predicting how physical quantities scale near the critical point, enabling a deeper understanding of universal aspects of critical behaviour (Fisher, 1998). Universality is a key concept, asserting that systems with distinct microscopic interactions can exhibit the same critical behaviour if they belong to the same universality class. This concept has revolutionised the study of critical phenomena, showing that diverse systems such as magnetic materials, fluid systems, and percolation models can exhibit analogous critical behaviours (Goldenfeld, 1992; Cardy, 1996).

1.5 Order Parameters

Order parameters are fundamental quantities that characterise the degree of order in a system and provide a clear diagnostic for identifying different phases of matter. As a system undergoes a phase transition, the order parameter typically changes either continuously or discontinuously, reflecting the symmetry-breaking or restoration that accompanies the transition. For example, in ferromagnetic systems, the order parameter is the magnetisation, which measures the degree of alignment of magnetic moments. Above the Curie temperature, the magnetisation drops to zero, signifying the transition to a disordered paramagnetic phase, while below the Curie temperature, it becomes non-zero, indicating long-range magnetic order (Landau & Lifshitz, 1980).

In liquid-gas transitions, the order parameter is often the density difference between the two phases, and as the system nears the critical point, this difference tends to zero, illustrating the continuous nature of second-order transitions (Stanley, 1971). Identifying the correct order parameter is critical to understanding the nature of the transition. In complex systems, this can be challenging, particularly near critical points where emergent phenomena defy traditional classifications. Traditionally, frameworks such as mean-field theory and renormalization group methods have been used to identify order parameters and describe collective behaviour (Kardar, 2007).

1.6 Machine Learning and Phase Transitions

In recent years, machine learning techniques have emerged as powerful tools for addressing the complexities of phase transitions. These techniques, particularly unsupervised learning algorithms, have shown promise in analysing

high-dimensional datasets and revealing hidden patterns that traditional methods might overlook. By identifying correlations and structures within the data, machine learning enables physicists to classify spin configurations in models like the Ising model, detect distinct phases, and study critical phenomena in a novel, data-driven manner.

Unsupervised learning techniques such as clustering and dimensionality reduction provide new perspectives on phase transitions, offering the ability to identify order parameters and uncover latent structures in the system. In particular, these methods have been applied to simplified models such as the Ising model, which serves as an ideal testbed due to its simplicity and rich theoretical foundation. The application of machine learning to phase transitions not only complements traditional methods but also opens up new avenues for exploring more complex systems (Carrasquilla & Melko, 2017; Wetzel, 2017).

2 Objectives

The primary objective of this dissertation is to employ unsupervised learning techniques to study phase transitions in simplified models of physical systems. Specifically, this work will focus on the 2D Ising model, a well-understood theoretical framework for studying phase behaviour and critical phenomena. By using computational simulations and machine learning analyses, this dissertation aims to uncover patterns in spin configurations that reveal the system's transition between ordered and disordered phases. Ultimately, this work seeks to validate the effectiveness of machine learning in capturing critical behaviour and provide new insights into phase transitions in complex systems.

2.1 Methodology

This research involves simulating the 2D Ising model across a range of temperatures using Monte Carlo methods, specifically the Metropolis-Hastings algorithm. These simulations will generate large datasets of spin configurations, capturing the system's evolution as temperature varies. Subsequently, unsupervised machine learning techniques—including clustering and dimensionality reduction—will be applied to these datasets to discern distinct phases and uncover latent structures. This integrated approach is designed to bridge traditional theoretical frameworks with modern machine learning tools, advancing our understanding of phase transitions.

3 Literature Review

3.1 Introduction

Phase transitions are fundamental phenomena in physics, characterised by abrupt changes in the properties of a system namely, the transformation of matter from one state to another under external influences such as temperature or pressure. Understanding these transitions is crucial for a wide range of scientific and technological applications. Phase transitions are critical in both theoretical and applied physics due to their role in understanding fundamental properties of systems. In condensed matter physics, they help explain phenomena such as superconductivity and superfluidity (Sachdev, 2011). In material science, phase transitions affect key material properties like magnetism and structural integrity, influencing the development of advanced materials (Anderson, 2018). Furthermore, in quantum computing, phase transitions can be harnessed to control quantum states, offering new approaches to

information storage and processing (Nielsen & Chuang, 2010). Thus, studying phase transitions has both fundamental and practical importance. In understanding phase transitions one of the most studied models is the 2D Ising Model, which provides critical insights into magnetisation and order-disorder transitions this is because it provides a clear framework for studying critical behaviour, spontaneous system breaking, and universality in both 2D and 3D systems. (Goldenfeld, 1992). Machine learning is revolutionising the analysis of phase transitions by addressing the limitations of traditional computational methods. Techniques like neural networks and unsupervised learning can efficiently identify critical points and phase transitions without requiring prior knowledge of the system's order parameters. Machine learning models can handle complex, high-dimensional systems where traditional methods, such as Monte Carlo simulations, struggle due to issues like critical slowing down (Carrasquilla & Melko, 2017). These models enable more accurate and scalable analysis of phase transitions. The aim of this review is to investigate the recent advancements in the use of machine learning techniques to detect phase transitions in the Ising model and highlight how data-drive approaches are revolutionising this area of study.

3.2 The Ising model

The Ising model, introduced by Ernst Ising in 1925, remains a cornerstone in the study of phase transitions due to its simplicity and analytical tractability. It describes a system of binary spins interacting with nearest neighbours on a lattice, making it ideal for exploring spontaneous magnetization and critical phenomena (Ising, 1925). The Ising model is widely used as a benchmark in machine learning applications because it captures the essence of phase transitions, providing a framework for testing algorithms aimed at identifying

critical points. For instance, Carrasquilla and Melko (2017) applied neural networks to the 2D Ising model to distinguish between ferromagnetic and paramagnetic phases, showing that machine learning models can efficiently detect the critical temperature without prior knowledge of the system's properties which allows it to overcome the challenges faced by more traditional methods such as monte Carlo simulations. (Carrasquilla & Melko, 2017)

The significance of the Ising model extends beyond its theoretical simplicity; its use in machine learning highlights the ability of data-driven approaches to learn from microscopic configurations and identify emergent macroscopic behaviour. Additionally, unsupervised learning methods such as Principal Component Analysis (PCA) and variational autoencoders have been used to explore the phase structure of the Ising model, showcasing the model's versatility in machine learning research (Wetzel, 2017). The Ising model thus serves as an essential testbed for developing and refining machine learning techniques that aim to detect phase transitions in more complex and high-dimensional systems.

3.3 Machine Learning in Physics

In recent years, the intersection of machine learning (ML) and physics has emerged as a powerful approach to studying complex physical systems, particularly in areas such as phase transitions. This new interdisciplinary field leverages the predictive capabilities and data-driven insights of ML to complement traditional physics methods, which often rely on analytical or computational models. For example, Schindler et al. applied unsupervised learning techniques to investigate many-body localization in quantum systems, using neural networks to classify different phases. Specifically, they employed clus-

tering algorithms to analyse spin configurations in the Ising model, showing that these methods could automatically identify critical points and classify phases without predefined order parameters. Their approach demonstrated the potential of unsupervised learning to detect hidden patterns in complex systems where traditional methods may fail, emphasising the power of machine learning in revealing new insights into phase transitions.

Neural networks have been successful in the detection of phases transitions. Carrasquilla and Melko (2017) were able to demonstrate that neural networks could automatically classify phases in the 2D Ising model without prior knowledge of the system's order parameters. However as successful as the results may look, neural networks lack interpretability due to their 'black box' nature which means that the humans who design the algorithm cannot understand how the variables are being combined to make predictions. (Carrasquilla & Melko, 2017). Neural networks automatically learn complex relationships from the data but do not provide clear explanations for how they make predictions or classifications. In this context, although the phases could be accurately identified it does not offer any physical insights into the underlying mechanisms or critical behaviours driving the transitions. Unlike traditional model or simpler machine learning models, neural networks do not provide direct access to interpretable parameters like order parameters or critical exponents; this poses an issue in physics, where understanding critical behaviour is as important as accurate phase detection. (Montavon, Samek & Müller, 2018). Montavon, Samek & Muller (2018) emphasise that while neural networks excel in classification tasks, their inability to offer physical interpretability limits their broader application in scientific research, where understanding system dynamics is crucial, in contrast, simpler models like SVMs provide more interpretable decision boundaries, but at the cost of

flexibility. This trade-off between accuracy and interpretability remains a key challenge in the world of applying machine learning to physical phenomena like phase transitions. (Montavon, Samek & Müller, 2018)

In their 2018 study, Morningstar and Melko explored the use of deep learning to analyse the 2D Ising model near criticality. They employed deep neural networks to classify phases of matter and identify the critical temperature. Notably, their work demonstrated that the neural networks could successfully detect phase transitions even without prior knowledge of the system's order parameters, effectively learning from the raw spin configurations. The networks were able to extract key features from the data, such as the magnetisation and susceptibility, which typically characterise phase transitions. Their approach highlighted the potential of deep learning to automate the analysis of critical phenomena, providing both accuracy and efficiency, especially near critical points where traditional methods are computationally expensive. (Morningstar & Melko, 2018) However again this study brought to the forefront the common challenge of neural networks for physics there was a lack of interpretability; while the networks could predict the critical temperature and classify the phase transitions, they provided limited insight into the physics that drives these transitions. (Montavon, Samek & Müller, 2018)

Rodriguez-Nieva and Scheurer (2019) applied convolutional neural networks (CNNs) to identify topological phase transitions, focusing on systems where traditional methods struggle due to the complexity of the interactions. They demonstrated that CNNs could detect critical behaviour and topological features by analysing spatial correlations in the system's data, outperforming conventional approaches. Their work showed that machine learning models,

particularly CNNs, are well-suited for identifying spatially complex transitions, such as those present in high-dimensional or topologically ordered systems. CNNs excel at identifying spatial correlations in high-dimensional systems, but they require large datasets and substantial computational power for training, making them impractical for certain applications. (Rodriguez-Nieva & Scheurer, 2019)

Ponte and Melko (2017) also applied support vector machines (SVMs) to the study of phase transitions. They were used to classify phases in the ising model to predict the critical temperature with a high degree of accuracy. SVMs are advantageous for systems with high-dimensional data, allowing them to efficiently find the decisions boundaries that separate different phases. This study highlights the effectiveness of SVMs in the detection of phases transitions and the interpretability of this machine learning tool unlike neural networks. (Ponte & Melko, 2017) However, SVMs are less flexible than neural networks meaning that SVMs are less likely to automatically learn intricate patterns or features from data whereas neural networks, particularly deep learning models, can automatically adjust their internal representations and handle more complex relationships due to their layered architecture, making them more adaptable to a wider range of problems. The decreased flexibility of SVMs is because they rely on predefined kernel functions to map data into higher dimensions which limits their ability to learn more complex relationships. (Goodfellow, Bengio & Courville, 2016)

Other unsupervised learning techniques such as clustering techniques and principal component analysis (PCA) have risen as powerful tools in the detection of phase transitions. Wetzel (2017) showed that PCA could be used to distinguish phases in the Ising model and detect critical points by detecting

patterns in spin configurations. PCA is an efficient method and requires less data in comparison to neural networks however, again it is limited by its linear nature and may miss intricate non-linear relationships present within the system. Wetzel's study highlights that while PCA highlights interpretable and computationally inexpensive tool, it struggles to capture the rich critical behaviours seen in higher dimensional quantum systems. In Wetzel's study he used other non-linear techniques such as autoencoders to capture the more complex relationships in phase transitions. (Wetzel, 2017)

Applications of machine learning extend beyond well-known models like the Ising model. It has been used to identify complex phases in systems such as spin glasses, quantum many-body systems, and high-dimensional materials, where traditional methods struggle due to the sheer number of degrees of freedom and the complexity of interactions (Wetzel, 2017). Unsupervised learning techniques have been used to discover hidden phases and phase transitions, revealing new insights into the behaviour of materials under various conditions. However, challenges remain, particularly in ensuring that machine learning models respect the underlying physics of the system and are interpretable by physicists. There is ongoing research aimed at integrating machine learning more deeply with physical principles, with promising future directions including hybrid approaches that combine data-driven methods with theoretical models to improve accuracy and provide deeper insights into phase transitions (Carleo et al., 2019).

Other recent studies have applied machine learning techniques to detect phase transitions in quantum materials and spin glass systems, areas where traditional methods face significant challenges in doing this. Van Nieuwenburg et al. (2017) introduced the "Confusion Method", where neural networks were

trained to classify phases in quantum systems by generating artificial data. This approach allowed the networks to identify critical points in systems with no explicit order parameter, highlighting the ability of machine learning to uncover hidden quantum phases. (Van Nieuwenburg, Liu & Huber, 2017)

Morningstar et al. (2021) applied deep learning to the study of spin glass systems, which are notoriously difficult to analyse due to their rugged landscapes. Traditional methods often struggle with complex local minima and the disordered nature of spin glasses, but the authors showed that deep learning could effectively classify the glassy phases and capture the intricate correlations between spins. This work demonstrates how machine learning models can navigate high-dimensional spaces and provide insights into disordered systems, which are often computationally prohibitive using classical techniques. (Morningstar, Melko & Carrasquilla, 2021)

Moreover, machine learning has shown great promise in studying high-dimensional quantum materials. These systems, with their vast configuration spaces, often present formidable challenges for traditional Monte Carlo simulations and renormalisation group techniques. Machine learning methods, particularly neural networks and unsupervised learning models, offer a scalable alternative for analysing these complex systems. By extracting meaningful features from raw data, these models can efficiently classify phases, even in higher dimensions, where manual identification would be infeasible. These recent applications highlight the transformative role of machine learning in advancing our understanding of phase transitions across a broad range of physical systems. (Carleo et al., 2019; Van Nieuwenburg, Liu & Huber, 2017; Morningstar, Melko & Carrasquilla, 2021)

Conclusion & Future Research Directions

The integration of machine learning techniques with traditional methods is transforming the study of phase transitions, particularly in the Ising model. Machine learning models such as neural networks (Carrasquilla & Melko, 2017), SVMs (Ponte & Melko, 2017), and unsupervised learning techniques like PCA (Wetzel, 2017) have shown remarkable efficiency in detecting phase transitions and critical points. However, challenges remain in balancing interpretability with performance, as neural networks offer the greatest flexibility but lack interpretability; SVMs provide better interpretability but are limited in their complexity; and PCA excels in simplicity and computational efficiency but struggles with non-linear systems. (Montavon, Samek & Müller, 2018) The choice of method should depend on the specific demands of the phase transition detection task—whether the focus is on accuracy, interpretability, or computational efficiency.

Future research should aim to combine data-driven approaches with physical principles, enhancing both accuracy and insight into phase transitions. For example, to address the interpretability issues of neural networks in phase transition studies, future research could explore hybrid models that combine the strengths of both neural networks and more interpretable methods like SVMs. For instance, incorporating explainability techniques such as Layer-wise Relevance Propagation (LRP) could help physicists understand how neural networks make predictions at each layer, shedding light on critical behaviours (Montavon, Samek & Müller, 2018). Additionally, hybrid models could integrate neural networks with PCA, allowing the network to handle complex non-linear relationships while leveraging PCA's interpretability to highlight key phase features. Another promising avenue is reinforcement

learning, where agents could be trained to autonomously explore phase space and identify critical points, as demonstrated in early studies. Also, quantum machine learning has potential to revolutionise the field, with algorithms like quantum neural networks offering increased efficiency in processing large datasets in complex quantum systems (Biamonte et al., 2017). Combining these techniques with traditional methods could offer both enhanced performance and deeper insights into phase transitions.

4 Methodology

4.1 2D Ising Model

The 2D Ising model, introduced by Ernst Ising in 1925, represents a seminal contribution to statistical mechanics and the study of phase transitions. Originally formulated to understand ferromagnetism, the Ising model has become a cornerstone in the study of critical phenomena due to its simplicity and rich behaviour (Ising, 1925). The 2D Ising model is defined on a square lattice where each site is occupied by a spin variable that can take on values of +1 or -1, representing two possible states (e.g., spin-up or spin-down). Spins interact with their nearest neighbours. The energy if the system is determined by the alignment of the spins that neighbour each other: if two neighbouring spins are aligned the energy will be minimised, representing a more stable state (both spins being +1 or -1) and if the spins are not aligned the energy of the system is higher, representing a less stable configuration. (Stanley, 1971) The Hamiltonian of the system is given by:

$$H \;=\; -J \sum_{\langle i,j \rangle} S_i S_j \;-\; h \sum_i S_i$$

where J is the interaction strength between neighbouring spins, $\langle i, j \rangle$ denotes sum of nearest-neighbour pairs, S_i represents the spin at site i, S_j represents the spin at site j , and h is an external magnetic field (Ising, 1925; Kandel & Kaplan, 2004). The interaction term $-JS_iS_j$ encourages neighbouring spins to align, leading to either a ferromagnetic (if $J > 0$) or an antiferromagnetic (if $J < 0$) interaction. (Yeomans, 1992)

The temperature plays a crucial role in this model, as it introduces thermal fluctuations. At high temperatures, the thermal energy allows the spins to flip more easily, causing them to become disordered. As the temperature is lowered, spins tend to align more, resulting in an ordered state where most spins are in the same direction. (Newman & Barkema, 1999)

The main goal of this project is to simulate the Ising model across a range of temperatures using Monte Carlo techniques, specifically the Metropolis-Hastings algorithm. By doing so, I aim to observe the phase transition at the critical temperature and investigate how key physical quantities such as magnetisation and energy vary with temperature. Additionally, I will apply machine learning methods, including principal component analysis (PCA), K-means clustering, and t-SNE, to classify and visualise the different phases of the system.

4.2 Monte Carlo Methods

Introduction to Monte Carlo Simulations

Monte Carlo methods are widely used in physics for simulating complex systems because they provide a probabilistic approach to solving problems that are often analytically unmanageable. Regarding the Ising model, the

Monte Carlo method allows for an efficient simulation of the system's evolution over time, particularly in systems which many components that interact, such as spins on a lattice. These methods are powerful in statistical physics because they enable the exploration of many possible configurations of the system at different temperatures, allowing us to observe how the system behaves statistically. (Newman & Barkema, 1999)

For the Ising model, the goal is to simulate the behaviour of spins on a lattice across a range of temperatures. Each spin interacts with its nearest neighbours, and the system tends to evolve toward configurations that minimise energy. The Monte Carlo method, particularly when combined with the Metropolis-Hastings algorithm, allows us to simulate this evolution efficiently by flipping spins randomly and evaluating the resulting change in energy. (Landau & Binder, 2021)

4.3 Metropolis Algorithm

The Metropolis-Hastings algorithm is a widely used method for generating configurations in Monte Carlo simulations, particularly for systems like the 2D Ising model. It allows for the efficient exploration of a system's configuration space by introducing randomness into the selection of configurations, while ensuring that the system evolves towards equilibrium in a statistically sound manner (Metropolis et al., 1953). The algorithm is especially useful in situations where direct analytical solutions are difficult or impossible to obtain, such as in complex interacting systems of spins.

In the context of the Ising model, the Metropolis-Hastings algorithm works by proposing spin flips at random lattice sites and accepting or rejecting these flips based on the resulting change in energy and temperature. Below is a

detailed outline of the steps involved in the algorithm:

Steps of the Metropolis-Hastings Algorithm:

1. **Random Spin Selection**:

   At each step, a spin from the lattice is randomly selected. This ensures that each spin has an equal chance of being flipped, allowing the system to explore different configurations over time. The randomness of the selection is crucial to avoid bias in the simulation (Landau & Binder, 2021).

2. **Energy Calculation**:

   After selecting a spin, the change in energy ΔE resulting from flipping that spin is calculated. The energy is determined by the interaction between the selected spin and its nearest neighbours, according to the Ising model Hamiltonian:

$$E = -J \sum_{\langle i,j \rangle} s_i s_j$$

Here, J is the interaction constant, and s_i and s_j represent the states of neighbouring spins (Newman & Barkema, 1999). The energy difference ΔE is computed by comparing the energy before and after the flip.

3. **Acceptance Criterion**:

   The spin flip is accepted or rejected based on the Metropolis acceptance criterion. If flipping the spin results in a lower energy (i.e., $\Delta E < 0$), the flip is always accepted, as it drives the system towards a more stable configuration. However, if the energy increases (i.e., $\Delta E > 0$ the flip is

accepted with a probability given by the Boltzmann factor:

$$P = e^{-\frac{\Delta E}{k_B T}}$$

where T is the temperature, k_B is the Boltzmann constant, and ΔE is the energy difference. This probability allows the system to occasionally accept higher-energy configurations, which prevents it from getting trapped in local energy minima and encourages proper exploration of the configuration space (Krauth, 2006).

4. **Periodic Boundary Conditions**:
   To avoid edge effects that could distort the simulation, periodic boundary conditions are applied. This means that spins on the edges of the lattice interact with spins on the opposite side of the lattice, effectively creating a system without boundaries. This technique ensures that all spins are treated equally and prevents artificial boundaries from influencing the system's behaviour (Binder & Heermann, 2010).

5. **Iteration and Equilibration**:
   This process of randomly selecting spins, calculating energy differences, and accepting or rejecting spin flips is repeated many times. Over the course of the simulation, the system evolves toward thermal equilibrium, where the physical quantities (such as energy and magnetisation) fluctuate around steady average values. The number of iterations required to reach equilibrium depends on factors like the size of the system and the temperature.

By iterating through these steps, the Metropolis-Hastings algorithm allows

for an accurate sampling of the spin configurations of the Ising model at different temperatures. This enables the observation of phase transitions and critical phenomena as the system shifts from an ordered (low-temperature) to a disordered (high-temperature) state (Stanley, 1971).

4.4 Simulation Parameters

When simulating the Ising model, several key parameters must be defined to ensure the system behaves realistically and provides meaningful insights into phase transitions. These parameters include the grid size, temperature range, number of Monte Carlo steps, and the method for determining when the system reaches equilibrium. Proper selection and handling of these parameters are essential for accurately modelling the system's behaviour.

4.4.1 Grid Size

For this simulation, a 32x32 lattice was chosen to represent the 2D grid of spins. Each point on the grid corresponds to a spin that can take values of either +1 or -1. The grid size is a compromise between computational efficiency and ensuring that the system is large enough to exhibit collective behaviour, especially near the critical temperature T_c . Smaller grid sizes can lead to finite-size effects, where phase transitions are less sharply defined due to the small number of interacting spins (Landau & Binder, 2021). A grid size of 32x32 ensures that the system is large enough to capture phase transitions while maintaining computational tractability.

4.5 Calculation of Physical Quantities

4.5.1 Energy

The total energy of the system is calculated based on the interactions between neighbouring spins. In the Ising model, each spin interacts only with its nearest neighbours, and the system's total energy E is given by the Hamiltonian:

$$E = -J \sum_{\langle i,j \rangle} s_i s_j$$

where s_i and s_j represent the spin values (either +1 or -1) at neighbouring sites i and j, and J is the coupling constant, which is positive for ferromagnetic interactions. The sum is taken over all nearest-neighbour pairs on the lattice (Newman & Barkema, 1999).

For each spin, the product of the spin at that site and the spins at its four neighbouring sites (left, right, above, and below, accounting for periodic boundary conditions) is calculated. The energy contribution from each pair of neighbours is added up to obtain the total energy of the system. Because each pair of neighbours is counted twice in this process, the final energy is divided by 2 to avoid double-counting:

$$E = \frac{-J}{2} \sum_{\langle i,j \rangle} s_i s_j$$

The energy gives us insight into the stability of the system. At lower temperatures, where spins tend to align, the energy is lower due to the favourable interactions between aligned spins. At higher temperatures, the system becomes disordered, and the energy increases as more spins become misaligned

with their neighbours.

4.5.2 Magnetisation

The magnetisation M of the system measures the degree of spin alignment across the entire lattice. It is calculated as the sum of all spins in the system:

$$M = \sum_{I} s_i$$

where s_i is the spin at site i on the lattice (**Landau & Binder, 2021**). Magnetisation reflects the net magnetic moment of the system: if most spins are aligned, the magnetisation will be high (close to ± 1), indicating an ordered phase. In a disordered phase, where spins are randomly oriented, the magnetisation approaches zero.

4.6 Machine Learning Analysis

4.7 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) was employed in this project as a dimensionality reduction technique to capture the most significant patterns in the spin configurations of the 2D Ising model. Since each spin configuration consists of many interacting spins (e.g., a 32x32 lattice has 1024 spins), the data is high-dimensional. By reducing the complexity of the dataset, PCA allows for easier visualisation and analysis while still retaining the most important information. Specifically, PCA helps identify the dominant features that explain the variance in spin configurations across different temperatures. This reduction in dimensionality is essential for simplifying the data while maintaining its underlying structure (Jolliffe & Cadima, 2016).

PCA works by projecting the high-dimensional data (the spin configurations) onto a lower-dimensional subspace. Each principal component represents a linear combination of the original features (spins) and captures the directions of maximum variance in the data. The principal components are ordered by the amount of variance they explain, with the first component capturing the most variance, the second capturing the next most variance, and so on.

Mathematically, PCA performs an eigenvalue decomposition of the covariance matrix of the data, finding the eigenvectors (the principal components) and eigenvalues (the amount of variance explained by each component). This technique reduces the data's dimensionality by projecting it onto a smaller number of components while retaining the most important information. Typically, the first few principal components explain most of the variance in the dataset, making it possible to ignore the rest without significant loss of information (Jolliffe, 2002).

The number of principal components to retain is chosen based on the explained variance. In this project, the first 7 components were retained because they captured a significant portion of the variance in the dataset, often explaining over 90% of the total variance. This reduction in dimensionality makes it easier to visualise and analyse the data, especially when looking for patterns related to phase transitions in the Ising model.

For example, by retaining only the first two components, the spin configurations in 2D space can be visualised, which allows us to easily observe clustering behaviour and phase transitions across different temperature regimes. In many cases, these two components are sufficient to reveal the underlying structure of the data (Jolliffe & Cadima, 2016).

The application of PCA resulted in the projection of the high-dimensional spin configurations onto a lower-dimensional space (e.g., 2D). Similar configurations—such as those from the same temperature range—tended to cluster together in this reduced space. This clustering behaviour provided insight into how the system transitions between phases. For instance, spin configurations at low temperatures (where the system is ordered) formed distinct clusters separate from those at high temperatures (where the system is disordered). The use of PCA also revealed the complexity of the spin configurations near the critical temperature T_c, where fluctuations in spin states were more pronounced and the data showed more spread in the lower-dimensional space (Bishop, 2006).

By reducing the dimensionality while retaining the key patterns, PCA was an effective tool for visualising and analysing the phase transition behaviour in the Ising model. It simplified the complexity of the dataset, making it easier to apply further machine learning techniques such as K-means clustering and t-SNE.

4.8 K-means Clustering

The K-means algorithm is a widely used unsupervised learning technique that divides data into a predefined number of clusters based on similarity. In this project, K-means was used to classify the spin configurations of the 2D Ising model at different temperatures into distinct clusters, which correspond to different phases of the system. By grouping spin configurations with similar patterns, K-means provided insight into how the system transitions between phases as the temperature changes. Specifically, the algorithm helped identify two key phases: the ordered phase (low temperature) and the disordered phase (high temperature) (MacQueen, 1967; Bishop, 2006).

In this project, the value of $K = 2$ was chosen for K-means based on the hypothesis that the Ising model undergoes a phase transition between two primary phases: the ordered phase (ferromagnetic) at low temperatures and the disordered phase (paramagnetic) at high temperatures. This binary clustering assumption is supported by the fact that the 2D Ising model is known to exhibit such a transition. Therefore, setting $K = 2$ allowed K-means to divide the data into two clusters that corresponded to these two distinct physical phases (Newman & Barkema, 1999).

The K-means algorithm works iteratively by performing the following steps:

1. **Cluster Initialisation**: The algorithm begins by randomly selecting K initial cluster centres.

2. **Assignment Step**: Each spin configuration (or its reduced representation) is assigned to the nearest cluster centre based on **Euclidean distance**.

3. **Update Step**: The cluster centres are then updated by calculating the mean of the configurations assigned to each cluster.

4. **Convergence**: These steps are repeated until the cluster assignments no longer change, meaning the algorithm has converged.

By grouping configurations into two clusters, K-means classified the configurations into ordered and disordered phases. The Euclidean distance metric ensured that similar configurations, such as those from the same temperature range, were grouped together, while different configurations were separated into distinct clusters (MacQueen, 1967).

The application of K-means clustering resulted in a clear distinction between the ordered and disordered phases of the 2D Ising model. Configurations at low temperatures, where the system exhibits high alignment of spins, were grouped into one cluster, corresponding to the ordered phase. Configurations at high temperatures, where the spins are more randomly oriented, were grouped into the second cluster, representing the disordered phase.

The result of K-means clustering confirmed the phase transition behaviour, where the critical temperature T_c marks the boundary between these two phases. The algorithm's ability to separate the data into these two clusters provided a clear and interpretable classification of the system's different physical phases (Lloyd, 1982; Bishop, 2006).

4.9 t-SNE

t-distributed Stochastic Neighbour Embedding (t-SNE) is a dimensionality reduction technique commonly used for visualising high-dimensional datasets. Unlike PCA, which focuses on preserving global variance, t-SNE prioritises preserving local structure in the data, making it an ideal method for identifying clusters or patterns in complex datasets. By projecting high-dimensional data into a lower-dimensional space (usually 2D or 3D), t-SNE makes it easier to see how similar data points are grouped, even in nonlinear datasets (van der Maaten & Hinton, 2008).

In this project, t-SNE was used to visualise the spin configurations of the Ising model across different temperatures, allowing us to observe patterns such as phase transitions in the system.

Parameters:

To generate a 2D visualisation of the data, t-SNE was run with the following parameters:

- **n_components=2**: This parameter specifies the number of dimensions for the output. In this case, 2D was chosen to allow for easy visualisation of the clusters.

- **perplexity=30**: Perplexity controls the balance between local and global data relationships. A value of 30 was chosen as it is a common default that works well for many datasets. It reflects the effective number of neighbours considered when calculating distances between points.

- **random_state=42**: A fixed random state was used to ensure that the results are reproducible.

These parameters were selected to provide a clear visualisation of the spin configurations and how they are grouped across different temperatures.

The t-SNE visualisation revealed distinct clusters that corresponded to different phases of the Ising model. Configurations from low temperatures, where spins are aligned (ordered phase), formed one cluster, while configurations from high temperatures, where spins are disordered, formed another. This clear separation of clusters confirmed that the system undergoes a phase transition as the temperature increases.

After performing K-means clustering, t-SNE was employed to visualise the clusters identified by K-means. While K-means provided a quantitative classification of spin configurations into ordered and disordered phases, t-SNE

allowed for a more intuitive, visual exploration of how these configurations are distributed.

In this approach, after running K-means on the dataset of spin configurations, t-SNE was applied to the same dataset, reducing it to a 2D visualisation space. The clusters formed by K-means were then colour-coded in the t-SNE plot to visualise how well the algorithm separated the phases. This helped confirm whether K-means accurately captured the boundary between the ordered and disordered phases of the 2D Ising model, particularly near the critical temperature T_c (van der Maaten & Hinton, 2008).

By combining t-SNE with K-means, I was able to both classify the spin configurations into phases (using K-means) and visualise these phases in a low-dimensional space (using t-SNE). This combined approach provided an intuitive way to understand the system's behaviour and the sharp transition between the ordered and disordered phases near the critical temperature T_c.

5 Results

The results presented in this section aim to demonstrate the phase transition in the 2D Ising model. By examining the system's evolution across a range of temperatures, we aim to identify key physical quantities, such as energy and magnetisation, that signal the transition from an ordered to a disordered phase. Additionally, the application of machine learning methods is intended to uncover latent structures in the spin configurations and verify whether they can accurately reflect the phase transition. This multi-faceted approach

allows us to validate both the theoretical predictions of phase transitions and the effectiveness of machine learning in capturing critical behaviour.

To achieve this, several techniques were applied, as outlined in the methodology. First, the physical quantities of energy and magnetisation are examined to provide a direct view of how the system's macroscopic properties evolve with temperature. Next, Principal Component Analysis (PCA) is used as a dimensionality reduction technique to uncover underlying patterns in the spin configurations. Following PCA, K-means clustering is applied to classify the spin configurations into distinct phases. Finally, t-SNE visualisations are used to further illustrate the clustering of configurations across different temperatures and highlight the distinction between ordered and disordered phases. Through this combination of methods, the results offer a comprehensive understanding of the phase transition in the 2D Ising model and the critical behaviour around T_c.

5.1 Energy

In the ordered phase, when the temperature is below the critical temperature $T_c \approx 2.269$, the system's energy remains relatively low and stable. This is because, at lower temperatures, the thermal fluctuations are minimal, allowing the ferromagnetic interactions between neighbouring spins to dominate, resulting in most of the spins being aligned. The alignment of spins minimises the system's energy, reflecting a stable ferromagnetic phase. In this phase, the spins are highly correlated, meaning that the energy remains in a low, stable range since the system exists in a relatively uniform state with few spin flips, as predicted by the Ising model's Hamiltonian (Stanley, 1971). This behaviour is typical for second-order phase transitions, where the energy does

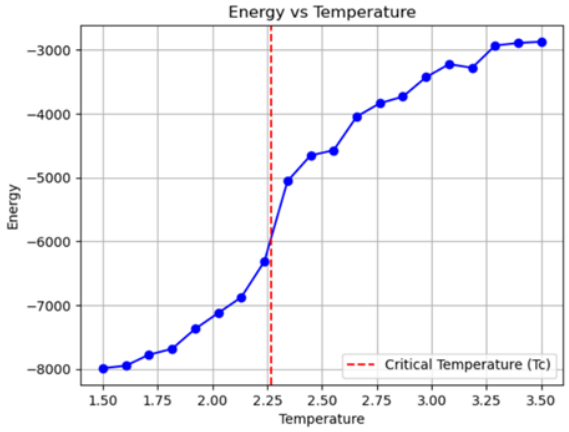

Figure 1: Energy as a function of temperature for the 2D Ising model. The plot shows the sharp rise in energy as the system approaches the critical temperature T_c, indicating the phase transition from the ordered (low-temperature) phase to the disordered (high-temperature) phase

not exhibit abrupt changes, unlike in first-order transitions where latent heat causes discontinuities (Goldenfeld, 1992).

As the system approaches the critical temperature $T_c$the total energy rises sharply, as seen in figure 1; this is because the thermal energy increases sufficiently to disrupt the alignment of spins, causing large-scale fluctuations. In this region, the system undergoes a phase transition, characterised by the breakdown of long-range order in the spin configurations. Near T_c, the ferromagnetic order is gradually destroyed as the spins begin to flip randomly due to thermal agitation, leading to increased energy. The sharp rise in energy around T_c is a direct manifestation of the critical phenomena associated with second-order phase transitions, where the correlation length diverges, and the system exhibits critical slowing down (Stanley, 1971).

This continuous change in energy is a hallmark of second-order phase tran-

sitions, in which there is no latent heat, but quantities like energy and magnetisation change continuously. The sharp rise near T_c represents the point at which the system transitions from an ordered ferromagnetic phase to a disordered paramagnetic phase (Kardar, 2007). The divergence in fluctuations at T_c reflects the growing instability as the system nears criticality, which is consistent with the theoretical framework of critical phenomena.

Once the system crosses $T_c$and enters the disordered phase at higher temperatures, the energy stabilises again, although at a higher value than in the ordered phase. In the disordered phase, thermal fluctuations dominate, and the spins become randomly oriented. At these temperatures, spin-spin interactions are weaker compared to the overwhelming effect of thermal energy, which causes the spins to randomly flip. As a result, the system settles into a higher, stable energy state, representing the paramagnetic phase where spins are uncorrelated and no longer exhibit long-range alignment. This stabilisation in the disordered phase is typical of second-order transitions, where the energy remains continuous, but now reflects the randomisation of spins (Newman & Barkema, 1999). The stabilisation of energy in the disordered phase is indicative of the system having reached a state of equilibrium, where random spin orientations no longer fluctuate significantly. The results here match well with the theoretical predictions of the Ising model, which suggest that energy in the paramagnetic phase fluctuates around a higher average value, driven by the dominance of thermal fluctuations over spin interactions (Yeomans, 1992).

The results align well with the theoretical understanding of second-order phase transitions. In the Ising model, a second-order phase transition occurs at T_c, marked by a smooth but sharp rise in energy as the system transitions

from the ordered to the disordered phase. This behaviour is consistent with the absence of latent heat, which differentiates second-order transitions from first-order transitions, where latent heat causes discontinuities in energy. In first-order transitions (e.g., liquid-gas transitions), the system absorbs or releases latent heat, leading to a sudden change in energy. However, in second-order transitions like the one observed here, energy changes continuously, as predicted by the renormalization group theory (Goldenfeld, 1992).

The energy behaviour in the simulation is consistent with the predictions of statistical mechanics and the Ising model, where thermal fluctuations increasingly disrupt spin alignment as temperature rises. This results in continuous but sharp changes in energy, confirming the system's critical behaviour at T_c. These findings provide strong evidence that the Ising model captures the essential features of real-world second-order phase transitions, reinforcing its importance in understanding critical phenomena (Stanley, 1971).

5.2 Magnetisation

At temperatures below the critical temperature T_c , the system is in the ordered phase where most of the spins are aligned. This strong alignment is reflected in a high, stable magnetisation value, as observed in the results. Magnetisation is a measure of the net magnetic moment of the system, which is proportional to the sum of spins in the lattice. When most of the spins are aligned in the same direction, the magnetisation approaches its maximum possible value. This behaviour is typical of a ferromagnetic phase, where the system minimises its energy by aligning spins, and thermal fluctuations are

not strong enough to disrupt this order (Stanley, 1971).

In this ordered phase, the magnetisation remains nearly constant because thermal energy is not yet sufficient to significantly flip the spins. This stability is expected in the Ising model, where low-temperature configurations are dominated by ferromagnetic interactions, resulting in a well-ordered state with strong magnetisation (Yeomans, 1992).

As the system approaches the critical temperature T_c , the magnetisation begins to drop sharply. This sharp decline is a hallmark of second-order phase transitions, where the order parameter (in this case, magnetisation) decreases continuously to zero at the critical point. The reduction in magnetisation reflects the growing influence of thermal fluctuations, which disrupt the spin alignment and weaken the long-range magnetic order. Near T_c , the system undergoes a phase transition from the ordered ferromagnetic phase to the disordered paramagnetic phase. This transition is driven by the competition between thermal energy, which tends to randomise the spins, and ferromagnetic interactions, which favour alignment (Kardar, 2007).

The sharp drop in magnetisation around T_c is consistent with the behaviour of second-order phase transitions, where the order parameter vanishes smoothly at the critical point. The Ising model predicts that at T_c , the system experiences large-scale spin fluctuations, resulting in a rapid loss of magnetisation. The correlation length, which measures the distance over which spins are correlated, diverges near T_c , leading to critical phenomena such as critical slowing down and enhanced fluctuations in spin configurations (Stanley, 1971). This explains the steep decline in magnetisation observed in the results, as the system transitions from an ordered to a disordered state.

Above the critical temperature, the system enters the disordered paramagnetic phase, where the magnetisation approaches zero. In this phase, the spins become randomly oriented due to the dominance of thermal fluctuations. The lack of spin alignment in the disordered phase means that the net magnetisation is close to zero, reflecting the fact that, on average, an equal number of spins point in opposite directions, cancelling out the overall magnetic moment (Newman & Barkema, 1999).

This behaviour is indicative of a system that has lost long-range magnetic order. At high temperatures, the thermal energy overcomes the ferromagnetic interactions between neighbouring spins, leading to a state of random spin orientations. As a result, the magnetisation remains near zero, signalling the absence of a coherent magnetic state. This is characteristic of the paramagnetic phase, where spins fluctuate independently, and no net magnetisation is observed (Yeomans, 1992).

The results align well with the theoretical predictions of second-order phase transitions in the 2D Ising model. In such transitions, the order parameter (magnetisation) decreases continuously as the system approaches the critical temperature. Unlike first-order transitions, where the order parameter exhibits a discontinuity, second-order transitions are marked by smooth, continuous changes. The magnetisation's sharp decline near T_c and its approach to zero at higher temperatures confirm that the system undergoes a second-order phase transition, as predicted by the Ising model (Goldenfeld, 1992).

Furthermore, the results demonstrate critical slowing down near T_c , where the system's response to perturbations becomes slower, and spin configurations exhibit large-scale fluctuations. This behaviour is reflected in the sharp drop

in magnetisation and the divergence of the correlation length, which are key signatures of critical phenomena in second-order transitions (Stanley, 1971). The agreement between the simulation results and the theoretical framework of phase transitions highlights the utility of the Ising model in capturing the essential physics of magnetisation behaviour near T_c .

By providing this deeper analysis of magnetisation behaviour, we can more clearly see how the simulation results illustrate the fundamental features of second-order phase transitions. The continuous but sharp decrease in magnetisation as the system transitions from the ordered to the disordered phase, and the approach to zero magnetisation in the paramagnetic phase, all align with well-established theoretical predictions in the study of critical phenomena.

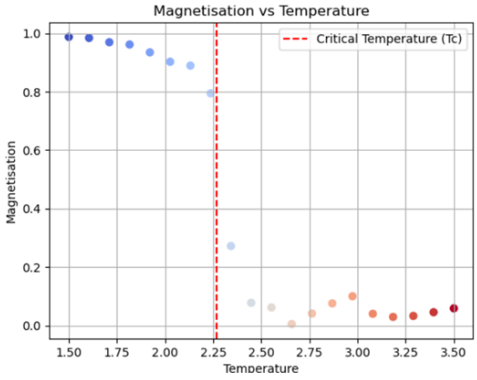

Figure 2: Magnetisation as a function of temperature for the 2D Ising model. The plot shows the sharp decline in magnetisation near the critical temperature T_c, indicating the phase transition from the ordered ferromagnetic phase at low temperatures to the disordered paramagnetic phase at higher temperatures.

5.3 PCA Analysis

Configurations below the critical temperature T_c cluster tightly together in the PCA space. This compact clustering suggests that the system's spin configurations are highly correlated, meaning that most of the spins are aligned, either pointing up or down. In the context of the Ising model, this corresponds to a state of low thermal agitation, where the spins minimise their energy by aligning due to strong ferromagnetic interactions (Stanley, 1971). The tight clustering in the PCA space indicates that there are relatively few distinct spin configurations at low temperatures, which is consistent with the expected physical behaviour of the system. In the ordered phase, where thermal fluctuations are weak, the system adopts a limited number of configurations, all of which are similar because the spins tend to align in the same direction. The low variance captured by PCA in this phase reflects the uniformity of spin alignments (Jolliffe & Cadima, 2016). This behaviour aligns with the theoretical predictions for ferromagnetic systems below T_c , where the magnetisation is high due to the collective alignment of spins. Because the system is in a low-energy state with strong correlations between neighbouring spins, the configurations in the ordered phase are highly similar, leading to a compact cluster in the PCA plot. The alignment of spins in the ordered phase minimises the system's energy, resulting in less variability across configurations.

In contrast, configurations from the disordered phase (above T_c) are more spread out in the PCA space, reflecting a greater diversity in spin configurations. This spread is caused by the thermal agitation at high temperatures, which disrupts the alignment of spins and leads to more random orientations. As thermal energy overcomes the ferromagnetic interactions, spins fluctuate

more freely, resulting in a wide range of possible configurations. This leads to higher variance in the PCA space, as the system explores a broader set of states (Stanley, 1971). The disordered phase is characterised by a loss of long-range order, where the spins no longer exhibit collective alignment. The broader spread of points in the PCA plot signifies the greater randomness in the system, as the correlation between neighbouring spins diminishes. In the disordered phase, thermal fluctuations dominate, and the system no longer maintains a stable magnetic moment. As a result, the spin configurations vary more widely, causing the PCA to show a more diffuse cluster (Bishop, 2006).

The first principal component (PC1) captures the largest portion of the variance in the data and likely corresponds to magnetisation, the key order parameter for the Ising model. This makes sense because magnetisation reflects the degree of spin alignment, which is the primary distinguishing feature between the ordered and disordered phases. In the ordered phase, where spins are aligned, the magnetisation is high, and the variance in configurations is low, leading to tight clustering along PC1. Conversely, in the disordered phase, magnetisation drops to near zero, and the variance in spin configurations increases, resulting in a more spread-out cluster in the PCA plot (Yeomans, 1992). This suggests that PC1 is strongly correlated with the overall alignment of spins, and thus with the system's magnetisation. In the ordered phase, where spins align, PC1 captures this high alignment and variance is low. In the disordered phase, PC1 captures the breakdown of this alignment as the system enters a more random state. The smooth transition of spin configurations from high to low magnetisation across the critical temperature reflects the continuous nature of second-order phase transitions (Stanley, 1971). The fact that PC1 explains most of the variance

also implies that magnetisation is the dominant feature in distinguishing the phases. Higher-order components, such as PC2, capture less variance and likely represent smaller-scale fluctuations within the phases, such as local spin flips that do not significantly affect the overall magnetisation. These components are more reflective of local behaviours rather than global phase transitions (Bishop, 2006).

The PCA results provide a clear visualisation of the phase transition behaviour in the 2D Ising model. The compact cluster of low-temperature configurations in the PCA plot signifies a high degree of spin alignment and order, while the more spread-out cluster of high-temperature configurations reflects the increasing disorder as the system crosses T_c. This transition is captured by the change in variance along PC1, which corresponds to magnetisation—the primary order parameter for the Ising model (Jolliffe & Cadima, 2016). As the system approaches T_c , thermal fluctuations become significant, leading to the rapid breakdown of spin alignment. The increase in variance along PC1 near the critical temperature is indicative of the critical phenomena that occur during second-order phase transitions, where the correlation length diverges, and the system exhibits large-scale fluctuations (Stanley, 1971). This behaviour is consistent with theoretical predictions, where the magnetisation drops continuously to zero at T_c, and the system transitions from a ferromagnetic to a paramagnetic state. The PCA results thus provide a clear, visual representation of the phase transition in the Ising model. By reducing the dimensionality of the spin configurations, PCA highlights the distinct clustering of configurations in the ordered and disordered phases, with PC1

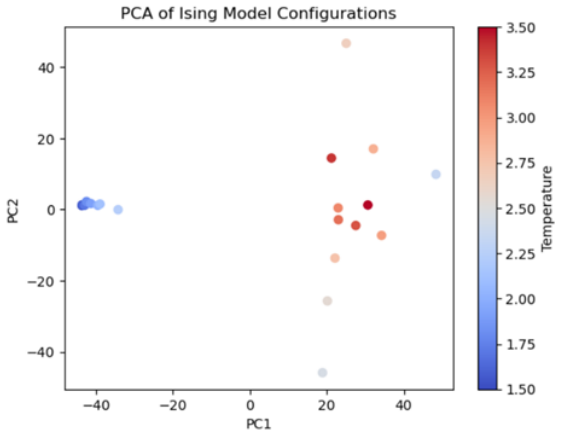

Figure 3: Principal Component Analysis (PCA) of the spin configurations of the 2D Ising model across different temperatures. The figure shows the clustering of low-temperature (ordered phase) and high-temperature (disordered phase) configurations, with the first principal component capturing most of the variance.

capturing the dominant feature—magnetisation—that drives this transition.

5.4 K-means Clustering

The K-means clustering plot shows two distinct clusters that correspond to the ordered (low temperature) and disordered (high temperature) phases of the 2D Ising model. This clustering behaviour aligns with the theoretical predictions of the Ising model, where the system undergoes a second-order phase transition at the critical temperature T_c . Below T_c, the system is in the ordered phase, characterised by a high degree of spin alignment, while above T_c, the system transitions to the disordered phase, where thermal fluctuations dominate, and spins become randomly oriented (Stanley, 1971).

In the ordered phase, the spins are highly aligned due to strong ferromagnetic

interactions, which minimise the system's energy. The tight clustering of configurations in this phase reflects the uniformity of spin alignments, meaning that the spin configurations are highly correlated. Most of the configurations in this cluster show similar spin arrangements, as the system tends to remain in a low-energy state with minimal thermal fluctuations at low temperatures. This is consistent with the physical behaviour of the Ising model below T_c , where the magnetisation is high, and the system exhibits long-range order (Yeomans, 1992). The high alignment of spins leads to a strong collective magnetisation, which is likely captured by the first principal component (PC1) in the PCA analysis. The K-means algorithm effectively groups these configurations into one cluster because they are similar, both in terms of spin alignment and energy.

Above the critical temperature T_c , the system enters the disordered phase, where thermal energy overcomes the ferromagnetic interactions, causing spins to fluctuate randomly. The broader spread of configurations in this phase, as seen in Cluster 2, reflects the higher variability in spin orientations. Unlike the ordered phase, where spin configurations are similar, the disordered phase is characterised by a lack of correlation between neighbouring spins. This randomness leads to greater diversity in spin configurations, which is why the K-means algorithm identifies this as a separate cluster. The system's magnetisation approaches zero in this phase, as the random orientation of spins cancels out any net magnetic moment. This behaviour is typical of the paramagnetic phase, where thermal fluctuations dominate, and the system no longer exhibits long-range order (Stanley, 1971).

The transition between these two phases occurs around the critical temperature T_c, where the system undergoes significant changes in its macroscopic

properties. At T_c, the correlation length diverges, and the system experiences large-scale fluctuations in spin configurations, which is captured by the PCA as a wider spread in configurations near T_c . The K-means algorithm, by separating the configurations into two clusters, provides a clear visual representation of the phase transition from order to disorder. The sharp distinction between the two clusters reflects the system's transition from a highly ordered state, where most spins align, to a disordered state, where thermal fluctuations lead to random spin orientations (Stanley, 1971).

The clustering results from K-means are consistent with the theoretical predictions for the 2D Ising model. In the ordered phase, spins are highly correlated, and the system remains in a low-energy, high-magnetisation state, leading to tight clustering. In the disordered phase, thermal fluctuations dominate, resulting in a greater diversity of spin configurations and a more spread-out cluster. The fact that K-means effectively separates these configurations into two distinct clusters highlights the model's ability to capture the transition between these phases. Additionally, the alignment of the K-means clusters with the principal components from PCA further validates that PC1 is likely capturing magnetisation, as it explains most of the variance between the ordered and disordered phases (Jolliffe & Cadima, 2016).

The results also imply that unsupervised learning methods like K-means clustering can be powerful tools in detecting phase transitions. By grouping configurations based on their similarities, K-means not only classifies the system into distinct phases but also provides insight into how the system evolves as temperature changes. The sharp separation between clusters highlights the dramatic change in the system's behaviour near T_c, where the system transitions from a magnetically ordered state to a disordered state.

This transition is a hallmark of second-order phase transitions, where the order parameter (in this case, magnetisation) changes continuously as the system crosses T_c, without the release of latent heat (Goldenfeld, 1992).

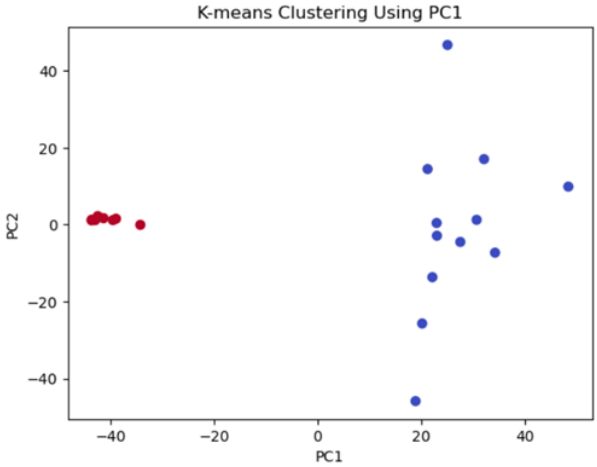

Figure 4: K-means clustering applied to the 2D Ising model spin configurations. The data are classified into two clusters representing the ordered (low-temperature) and disordered (high-temperature) phases, with the transition occurring near the critical temperature T_c.

5.5 t-SNE

The t-SNE plot, colour-coded by temperature, provides a powerful visualisation of how spin configurations evolve across different temperature regimes in the 2D Ising model. t-distributed Stochastic Neighbour Embedding (t-SNE) is a nonlinear dimensionality reduction technique that captures local similarities between data points while projecting them into a lower-dimensional space, typically 2D or 3D (van der Maaten & Hinton, 2008). By applying t-SNE to the spin configurations, we obtain a clear visual representation of the phase transition from the ordered to the disordered phase. In the t-SNE plot, low-temperature configurations (representing the ordered phase) are grouped

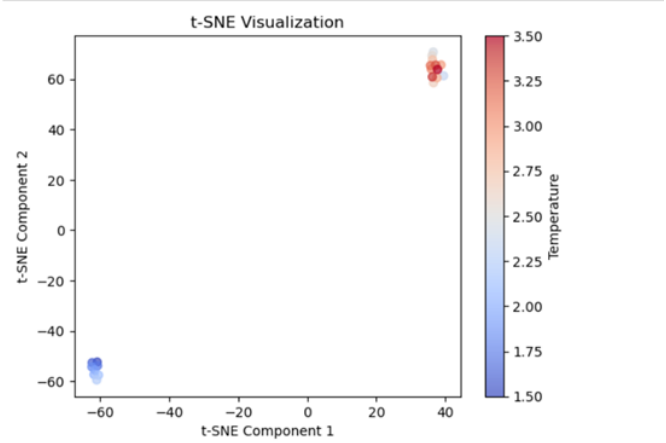

Figure 5: t-SNE visualisation of the 2D Ising model spin configurations across different temperatures. The plot highlights the separation between the ordered and disordered phases, showing distinct clusters for low- and high-temperature configurations, and a smooth transition near the critical temperature T_c.

tightly together, forming distinct clusters. These configurations share similar spin alignments, with most spins pointing in the same direction, which results in local similarities being preserved by t-SNE. High-temperature configurations (representing the disordered phase), however, are spread more broadly across the plot. This reflects the greater variability of spin orientations as thermal fluctuations increase at higher temperatures.

The t-SNE visualisation effectively separates the configurations from different temperature regimes, providing a clear distinction between the ordered and disordered phases. The clustering of low-temperature points and the broader distribution of high-temperature points highlight the system's transition from order to disorder as temperature increases. In the ordered phase, t-SNE groups configurations into compact clusters because spins are aligned, which

results in higher local similarity. This behaviour aligns with the physical expectation that in the ordered, ferromagnetic phase, spin configurations are highly correlated (Yeomans, 1992). As the system transitions into the disordered phase with rising temperature, thermal fluctuations lead to random spin orientations. t-SNE captures this by spreading out high-temperature configurations, indicating reduced similarity between them. The broader spread in this phase represents the system's greater configurational diversity as it loses long-range order (Stanley, 1971).

One of t-SNE's strengths is its ability to preserve local structure, meaning points close together in the plot have very similar original high-dimensional configurations. This feature makes t-SNE particularly well-suited for capturing how the spin configurations change incrementally as temperature increases, offering a smooth transition from the ordered to disordered phase.

To evaluate the clustering results, the Calinski-Harabasz index and the Silhouette score were employed. The Calinski-Harabasz index of 9651.00, a high value, indicates that the clusters are both well-separated and internally cohesive. This suggests a clear distinction between the ordered and disordered phases, reinforcing the robustness of the t-SNE visualisation. Additionally, the near-ideal Silhouette score of 0.97 confirms that points within each cluster are highly like one another, while points from different clusters are well-separated. Silhouette scores close to 1 are considered evidence of strong, well-defined clustering, further supporting the conclusion that t-SNE effectively distinguished between the two phases. The combination of these two metrics—both a high Calinski-Harabasz index and an almost perfect Silhouette score—validates that the clustering solution is robust and reliable, accurately representing the phase transition observed in the Ising model.

The results obtained through t-SNE, combined with those from PCA and K-means clustering, provide comprehensive insight into the phase transition in the 2D Ising model. The distinct clusters in the t-SNE visualisation illustrate the system's transition between the ferromagnetic and paramagnetic phases, and the analysis metrics provide quantitative evidence that t-SNE successfully captured these critical changes. This reinforces the idea that t-SNE is an effective tool for visualising and analysing phase transitions in complex systems.

5.6 Conclusion

In conclusion, the combined analysis of energy, magnetisation, PCA, K-means clustering, and t-SNE offers robust evidence of the phase transition in the 2D Ising model. The sharp changes in energy and magnetisation near T_c confirm the theoretical predictions of second-order phase transitions (Stanley, 1971). PCA effectively reduced the high-dimensional spin configurations into lower dimensions, revealing distinct clusters corresponding to physical phases (Jolliffe & Cadima, 2016). The application of K-means clustering and t-SNE further validated these results by providing a clear separation between the ordered and disordered phases, demonstrating the capability of machine learning techniques in detecting critical behaviour (van der Maaten & Hinton, 2008; MacQueen, 1967). Overall, the results align well with the theoretical understanding of phase transitions and offer strong evidence for the effectiveness of unsupervised learning methods in studying complex systems.

6 Discussion

The analysis conducted in this research has effectively demonstrated the utility of machine learning techniques in studying the phase transition of the 2D Ising model. The results obtained from the energy and magnetisation plots are consistent with the theoretical predictions of second-order phase transitions, confirming critical behaviour at the transition temperature T_c . This study aimed to apply unsupervised machine learning techniques to detect and characterise the phase transition. The results indicate that methods such as Principal Component Analysis (PCA), K-means clustering, and t-SNE can provide a clear and interpretable depiction of the transition from the ordered ferromagnetic phase to the disordered paramagnetic phase.

The energy versus temperature plot revealed a sharp increase in energy near the critical temperature T_c , aligning well with expectations for second-order phase transitions in the 2D Ising model. Below T_c, the system remains in a low-energy state due to the alignment of spins, consistent with ferromagnetic ordering (Yeomans, 1992). As the temperature increases, thermal fluctuations begin to disrupt this alignment, leading to a rapid rise in energy as the system enters the disordered phase. This sharp rise in energy around T_c reflects the breakdown of long-range spin alignment and increased thermal agitation, which is characteristic of second-order transitions where energy evolves continuously across the phase boundary (Stanley, 1971). Above T_c , the energy begins to stabilise again as the system reaches a disordered equilibrium state, where spins are randomly oriented, and the system's energy plateaus at a higher value (Newman & Barkema, 1999). Similarly, the magnetisation versus temperature plot further confirms the phase transition near T_c . Below T_c, the magnetisation remains high, reflecting strong spin

alignment in the ordered phase. However, as the system approaches the critical temperature, magnetisation decreases sharply, approaching zero as the system enters the paramagnetic disordered phase. This behaviour is indicative of second-order phase transitions, where the order parameter (in this case, magnetisation) continuously approaches zero as thermal fluctuations dominate and long-range order is lost (Stanley, 1971; Goldenfeld, 1992). The sharp drop in magnetisation signifies the breakdown of ferromagnetic ordering and the emergence of paramagnetism, where spins lose their alignment due to increasing thermal disorder. The system's critical slowing down near T_c , where large-scale fluctuations occur, marks this critical transition, and the results align well with theoretical predictions regarding how magnetisation behaves in such systems (Yeomans, 1992).

The application of machine learning techniques has proven instrumental in uncovering and visualising the patterns associated with the phase transition. PCA was particularly effective in reducing the dimensionality of the spin configuration data while retaining key features distinguishing the ordered and disordered phases. The clustering of low-temperature configurations into tight groups in the PCA plot reflects the homogeneity of the ordered phase, where spins are highly correlated due to ferromagnetic interactions. As temperature increases, PCA shows a broader spread of configurations in the disordered phase, where thermal fluctuations cause random spin orientations. The first principal component (PC1) captures the largest portion of the variance, suggesting that it is closely related to the system's order parameter, likely magnetisation. This aligns with the theoretical understanding that magnetisation is the key quantity driving the system's behaviour near the critical temperature (Jolliffe & Cadima, 2016). The ability of PCA to reduce high-dimensional data while preserving essential features makes it a powerful

tool for detecting phase transitions in the Ising model (Wetzel, 2017).

K-means clustering was also effective in classifying the spin configurations into two distinct clusters, corresponding to the ordered and disordered phases. The separation of the clusters in PCA space provides a clear representation of the phase transition, where configurations in the ordered phase group tightly together, and those in the disordered phase are more spread out. K-means was chosen for its simplicity and effectiveness in clustering high-dimensional data into clear categories based on the largest sources of variance (MacQueen, 1967). While K-means is a linear method, it captured the global structure of the system well by dividing configurations into distinct clusters that correspond to the phase transition near T_c . The clear alignment between the K-means results and the physical interpretation of the system's phases reinforces the value of using unsupervised learning methods to explore complex data in statistical mechanics.

t-SNE further enhanced the visualisation of the phase transition by highlighting local similarities between spin configurations. Unlike PCA, which focuses on capturing global variance, t-SNE preserves local structure, making it particularly effective for detecting subtle differences in configurations that arise near the phase transition. The t-SNE plot shows distinct clusters for the ordered and disordered phases, with configurations below T_c tightly grouped and those above T_c more dispersed. This result aligns with the physical expectation that spin configurations become more diverse as the system transitions from order to disorder (van der Maaten & Hinton, 2008). The high Silhouette score (0.97) and Calinski-Harabasz index (9651.00) provide quantitative validation of the quality of the clustering, confirming that the ordered and disordered phases are well-separated and internally cohesive (Rousseeuw,

1987; Calinski & Harabasz, 1974). The success of t-SNE in capturing the local structure of the data, particularly near the critical temperature, demonstrates the value of nonlinear dimensionality reduction techniques in studying phase transitions, where subtle local variations in the system's behaviour are critical.

While the current study demonstrates the effectiveness of PCA, K-means clustering, and t-SNE in detecting the phase transition in the 2D Ising model, several potential areas for future research remain unexplored. One key limitation of the present work is the finite system size, which tends to smooth out sharp transitions and obscure critical phenomena. Increasing the system size or extending the analysis to the 3D Ising model would provide more precise results. Recent studies (Landau & Binder, 2009) suggest that in larger systems, finite-size effects become less significant, allowing for a clearer understanding of the critical behaviour. Additionally, the unsupervised learning methods employed in this study, while effective, have their limitations. For instance, PCA and K-means clustering are linear methods and may struggle to capture complex, non-linear features of the phase transition near the critical temperature. In contrast, advanced machine learning techniques like Variational Autoencoders (VAEs) or Convolutional Neural Networks (CNNs) could offer deeper insights.

VAEs could be used to model the distribution of spin configurations, allowing the system to autonomously learn latent representations of the phase transition. By mapping high-dimensional data into a lower-dimensional latent space, VAEs could identify subtle patterns in the spin configurations that PCA and K-means might overlook. For example, Wetzel (2017) showed that VAEs can effectively capture non-linear correlations in the Ising model, making them a powerful tool for detecting critical points and identifying order parameters.

CNNs, commonly used in image processing, could also be applied to the spin configuration data. Since each lattice configuration can be treated as an image, CNNs can automatically learn spatial patterns and detect phase transitions by identifying distinct features in spin arrangements across temperature regimes. For instance, Carrasquilla and Melko (2017) demonstrated that CNNs can classify phases of matter in spin systems without any prior knowledge of the system's physical properties, offering a highly scalable approach for studying more complex models like the 3D Ising model.

Reinforcement learning (RL) is another promising avenue for future research. RL could be used to autonomously explore phase space and identify critical points with greater efficiency. Carleo et al. (2019) applied reinforcement learning to quantum many-body systems, showing that RL agents can be trained to detect phase transitions by maximising an objective function related to the system's thermodynamic properties. By employing RL in the study of the Ising model, future work could potentially reduce computational costs while improving accuracy in detecting transitions. Moreover, Explainability techniques such as Layer-wise Relevance Propagation (LRP) could be integrated with these advanced models to provide greater interpretability. One key challenge with neural networks is their "black box" nature, making it difficult to connect the learned features to physical quantities. LRP could be used to trace how decisions are made by machine learning models, enabling researchers to relate the extracted features back to the physical order parameters of the Ising model (Montavon, Samek & Müller, 2018). This would bridge the gap between the predictive power of machine learning models and their physical interpretability.

In summary, future work should explore advanced machine learning techniques

like VAEs, CNNs, and reinforcement learning to extend the analysis to larger or more complex systems, such as the 3D Ising model or quantum spin systems. Additionally, incorporating explainability techniques will ensure that these models remain physically interpretable, allowing us to draw more meaningful conclusions from their predictions.

The broader implications of this study underscore the transformative potential of machine learning (ML) beyond classical models like the 2D Ising model, extending into critical areas of physics such as quantum computing and materials science. The ability of ML algorithms to manage large-scale, high-dimensional datasets and uncover latent patterns makes them indispensable tools in fields where traditional analytical methods often face limitations. This has been especially evident in the study of quantum phase transitions and the development of quantum algorithms, where techniques such as variational autoencoders (VAEs) and reinforcement learning have shown promise in efficiently approximating quantum states and exploring vast quantum phase spaces (Carleo et al., 2019).

In quantum computing, ML has emerged as a key player in optimising quantum circuits and identifying optimal quantum states. The challenge of solving complex quantum many-body problems has been significantly addressed by ML techniques, which can approximate ground states and efficiently navigate through vast configuration spaces. Recent research highlights the potential of ML in developing quantum error correction algorithms, where the identification of error patterns is critical for the practical implementation of quantum computers (Melko et al., 2019). The ability to harness machine learning to control and stabilise quantum systems could accelerate the realisation of scalable quantum computing technologies. Similarly, in materials science, ML

has the potential to revolutionise the discovery of new materials with tailored properties. Through methods like convolutional neural networks (CNNs), ML models have been applied to predict phase diagrams and discover new compounds that meet specific performance criteria (Schmidt et al., 2019). These techniques offer a faster and more efficient alternative to traditional experimentation, which is often resource-intensive and time-consuming. For example, in the study of high-temperature superconductors and topological insulators, ML is being employed to predict material properties and explore phases of matter that are difficult to characterise using conventional approaches (Cubuk et al., 2019).

Moreover, ML's impact extends to complex systems like spin glasses and disordered materials, where the nature of the phase transitions can be intricate and hard to detect. Unsupervised learning techniques such as autoencoders and t-SNE have been shown to effectively identify hidden phases in these systems by revealing patterns that are otherwise challenging to uncover through traditional statistical mechanics alone (Rodriguez-Nieva & Scheurer, 2019). The ability of ML to handle non-linear relationships and uncover emergent behaviours in these systems reinforces its value in advancing our understanding of critical phenomena in highly complex environments.

In a broader context, the use of ML in physics represents a shift in how we approach scientific discovery. The integration of deep learning models with traditional physical methods opens new avenues for exploring quantum field theory, condensed matter physics, and cosmology, among other fields. For instance, in cosmology, ML has been used to process vast amounts of observational data to model the large-scale structure of the universe, while in condensed matter physics, it has helped classify topological phases that lack

conventional order parameters. The significance of ML in physics extends far beyond the Ising model. It is poised to become an essential tool in modern research, not just for simulating and modelling physical systems but for actively shaping the future of technologies like quantum computing and advanced materials. As machine learning continues to evolve, its application in these domains will likely lead to breakthroughs in areas previously considered too complex or computationally intractable, ultimately pushing the boundaries of both theoretical and experimental physics.

In conclusion, the results obtained in this research demonstrate the potential of machine learning techniques to detect and analyse phase transitions. The combination of traditional physical analysis with unsupervised learning methods has proven effective in capturing the phase behaviour of the 2D Ising model. Future work should explore more advanced algorithms, increase system size, and extend the analysis to higher-dimensional models to further understand the critical phenomena present in complex systems. As machine learning continues to evolve, its application in physics will likely lead to more sophisticated analyses of phase transitions and critical phenomena, advancing our understanding of these fundamental processes.

7 Reference List

1. Anderson, P. W. (2018). *Basic Notions of Condensed Matter Physics.* Addison-Wesley.

2. Atkins, P. (2010). *Physical Chemistry.* Oxford University Press.

3. Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N.,

& Lloyd, S. (2017). Quantum machine learning. *Nature*, 549(7671), 195–202.

4. Binder, K., & Heermann, D. W. (2010). *Monte Carlo Simulation in Statistical Physics.* Springer.

5. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* Springer.

6. Calinski, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods*, 3(1), 1–27.

7. Cardy, J. (1996). *Scaling and Renormalization in Statistical Physics.* Cambridge University Press.

8. Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., & Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4), 045002.

9. Carrasquilla, J., & Melko, R. G. (2017). Machine learning phases of matter. *Nature Physics*, 13(5), 431–434.

10. Cubuk, E. D., Sendek, A. D., Dwaraknath, S., Ding, H., Ramasubramaniam, A., & Reed, E. J. (2019). Machine learning and the search for new materials. *Nature Reviews Materials*, 4(1), 10-20.

11. Donnelly, R. J. (2009). *Quantized Vortices in Helium II.* Cambridge University Press.

12. Fisher, M. E. (1974). The renormalization group in the theory of critical behavior. *Reviews of Modern Physics*, 46(4), 597–616.

13. Fisher, M. E. (1998). Renormalization group theory: Its basis and formulation in statistical physics. *Reviews of Modern Physics*, 70(2), 653–681.

14. Gibbs, J. W. (1961). *Elementary Principles in Statistical Mechanics.* Dover Publications.

15. Goldenfeld, N. (1992). *Lectures on Phase Transitions and the Renormalization Group.* Addison-Wesley.

16. Goldenfeld, N. (2018). *Lectures on Phase Transitions and the Renormalization Group.* CRC Press.

17. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.

18. Huang, K. (1987). *Statistical Mechanics.* John Wiley & Sons.

19. Ising, E. (1925). Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1), 253-258.

20. Jolliffe, I. T. (2002). *Principal Component Analysis.* Springer.

21. Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.

22. Kandel, D., & Kaplan, T. A. (2004). Generalization of the Metropolis Algorithm. *Physical Review Letters*, 93(17), 170405.

23. Kardar, M. (2007). *Statistical Physics of Fields.* Cambridge University Press.

24. Kittel, C. (2005). *Introduction to Solid State Physics.* Wiley.

25. Krauth, W. (2006). *Statistical Mechanics: Algorithms and Computations.* Oxford University Press.

26. Landau, L. D., & Lifshitz, E. M. (1980). *Statistical Physics, Part 1.* Pergamon Press.

27. Landau, D. P., & Binder, K. (2021). *A Guide to Monte Carlo Simulations in Statistical Physics.* Cambridge University Press.

28. Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137.

29. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297). University of California Press.

30. Melko, R. G., Carleo, G., Carrasquilla, J., & Cirac, J. I. (2019). Restricted Boltzmann machines in quantum physics. *Nature Physics*, 15(8), 887–892.

31. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., &

Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21(6), 1087–1092.

32. Montavon, G., Samek, W., & Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73, 1–15.

33. Morningstar, A., & Melko, R. G. (2018). Deep learning the Ising model near criticality. *Journal of Machine Learning Research*, 18(1), 5975–5991.

34. Morningstar, A., Melko, R. G., & Carrasquilla, J. (2021). Deep learning spin glasses. *Nature Communications*, 12(1), 1101.

35. Newman, M. E. J., & Barkema, G. T. (1999). *Monte Carlo Methods in Statistical Physics.* Oxford University Press.

36. Nielsen, M. A., & Chuang, I. L. (2010). *Quantum Computation and Quantum Information.* Cambridge University Press.

37. Ponte, P., & Melko, R. G. (2017). Kernel methods for interpretable machine learning of order parameters. *Physical Review B*, 96(20), 205146.

38. Rodriguez-Nieva, J. F., & Scheurer, M. S. (2019). Identifying topological order through unsupervised machine learning. *Nature Physics*, 15(8), 790–795.

39. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and*

Applied Mathematics, 20, 53–65.

40. Sachdev, S. (2011). *Quantum Phase Transitions.* Cambridge University Press.

41. Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials*, 5(1), 83.

42. Sethna, J. P. (2006). *Statistical Mechanics: Entropy, Order Parameters, and Complexity.* Oxford University Press.

43. Stanley, H. E. (1971). *Introduction to Phase Transitions and Critical Phenomena.* Oxford University Press.

44. Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.

45. Van Nieuwenburg, E. P. L., Liu, Y. H., & Huber, S. D. (2017). Learning phase transitions by confusion. *Nature Physics*, 13(5), 435–439.

46. Wetzel, S. J. (2017). Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Physical Review E*, 96(2), 022140.

47. Yeomans, J. M. (1992). *Statistical Mechanics of Phase Transitions.* Oxford University Press.