IROP Report

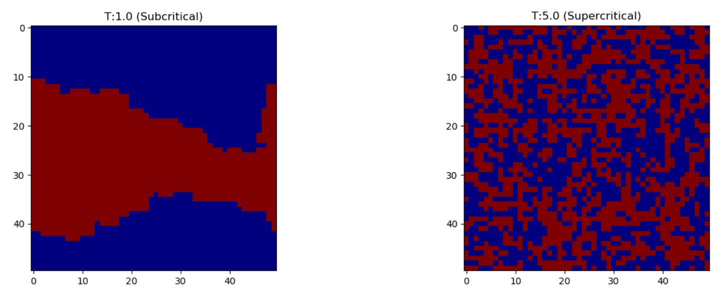
Robert King

September 2019

This summer I was fortunate enough to be selected to participate in Imperial's International Research Opportunities program (or IROP) where I had the chance to conduct research at the University of British Columbia in Vancouver, BC, Canada.

The focus of my research this summer was in applying Neural Networks to statistical models in physics, more specifically the Ising Model. The Ising Model is a simple model that describes ferromagnetic behaviour and is one of the simplest statistical models to display a phase transition. A phase transition is defined as a rapid change in the behaviour at a given temperature, known as the critical temperature or T_C. In the case of the Ising Model, the phase transition observed is between the ordered phase, where all the spins in the lattice are aligned in one direction and the disordered phase where the spin direction is entirely randomised. The first part of my project was to create a Monte Carlo simulation that replicated this behaviour. You can see below in figure 1 a comparison of a configuration taken at cold temperatures ($T << T_C$) and a configuration at a hot temperature ($T >> T_C$). As the temperature of the lattice increases the single large domain begins to split into many smaller domains of spins all aligned in the same direction until the critical temperature is reached. At that point the domains begin to break down into pure noise and there is no magnetic order at all left in the system. By plotting the net magnetization of the lattice, that is the sum of all the spin directions in the lattice and plotting this as a function of the temperature, the phase transition can be clearly seen. This is demonstrated in figure 2.

One interesting thing that is seen with the Ising Model is that snapshots of lattice configurations at temperatures slightly above or slightly below the critical temperature appear near indistinguishable. This leads on to the main focus of the project, to try and see if a neural network can detect and learn a difference between the two samples: one *sub-critical* ($T < T_C$) and the other *super-critical* ($T > T_C$). The first test was to use a so called "feed forward" neural network. In this structure the network is divided up into several "hidden layers" between the output and input neurons. These intermediate layers are fully connected to the previous and subsequent layers. This neural network architecture is able to distinguish temperature of ± 0.1 away from the critical temperature with greater than 99% accuracy. Around the critical temperature the performance begins to degrade however to some extent this is to be expected. From a statistical physics perspective it is known that any given lattice configuration can be valid across all temperatures. Configurations with energies far away than the statistical average however are statistically exceptionally rare. For two temperatures close together on either side of T_C, a configuration drawn at one temperature is also a likely to be a probable configuration of the other temperature. This is demonstrated in figure 3. The percentage "overlap" between the two distributions represents some permanent degree of uncertainty between the two temperatures and gives a "maximum" obtainable accuracy. By comparing this to the actual training accuracy of the network there is a gap of around 10% at temperature separations below ± 0.05. This gap represents that theoretical headroom



(a) (b)

Figure 1: Two configurations of the Ising Model. In a) the Temperature is much less than the critical temperature with large scale order seen, whilst in b) the temperature is much greater than T_C and the sample is near complete noise.

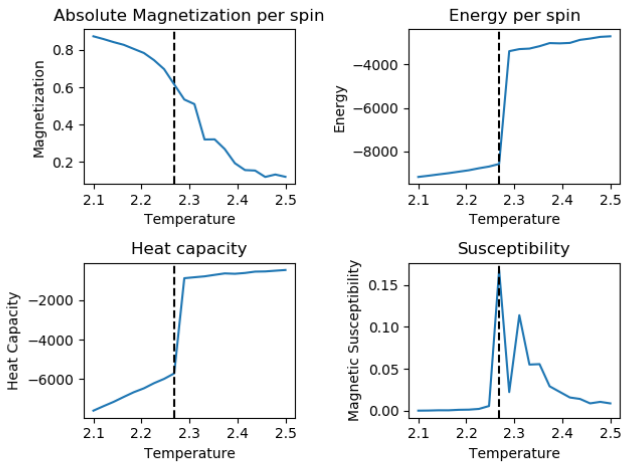


Figure 2: Plots showing the behaviour of various thermodynamic quantities as the temperature of the lattice is increased from $2.1k_B$ to $2.5k_B$. The dashed line indicates the critical temperature which is at $T_C \approx 2.27k_B$. The phase transition can be seen most notably in the absolute magnetization and in the magnetic susceptibility, where a sharp spike is seen near T_C.

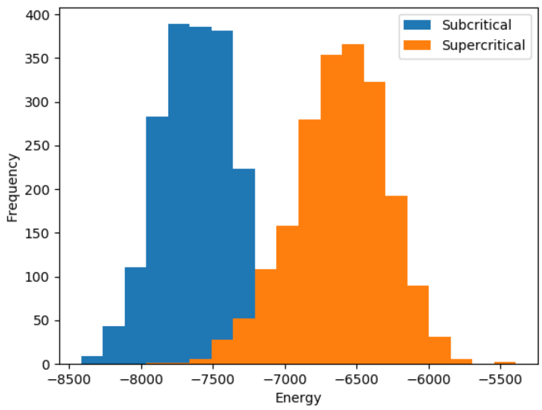


Figure 3: Plot showing the energy distribution of configurations above and below the critical temperature at $T_c \pm 0.05$. There is an overlap where at the intermediate energies, configurations were drawn from both the super and sub critical simulations. The percentage of configurations that were overlapping in this figure is 6.5% which implies that the best possible network performance is 93.5%.

that existed for network optimization.

One improvement that was made was to use convolutional neural networks (CNNs). These have the advantage of preserving the original structure and shapes present in the original 2D images. When using CNNs the training accuracy improved by around 7% from 80% to 87% still a fairway from the maximum obtainable accuracy of around 93%. Towards the end of the project the idea of using residual connections inside both network structures was investigated. This involves allowing connections in the network to allow for larger more complex networks to be tested. This helps to mitigate the problem of the input signal being lost due to as the network depth increased.

Overall this project was a really interesting investigation for me and gave me the opportunity to further develop my existing interest in Machine Learning methods and Computational Physics. The IROP program also gave me an amazing opportunity to explore Canada. I went into the program not knowing very much at all about Canada and British Columbia and now after the project I would definitely consider returning to Vancouver in the future!