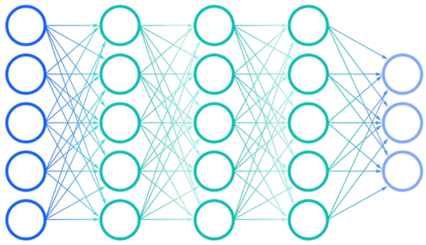
Science & Technology



Physics Meets Machine Learning

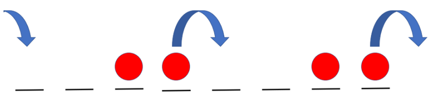
The physics of equilibrium systems is well known and can be solved by a one-size-fits-all approach. The same cannot be said for nonequilibrium systems which require more novel methods. A modern branch of datascience, machine learning (ML), has been applied to a decades old nonequilibrium physics model to show that ML tools deserve a place in a physicist's toolbox.

Nonequilibrium states with machine learning

by M. Stewart

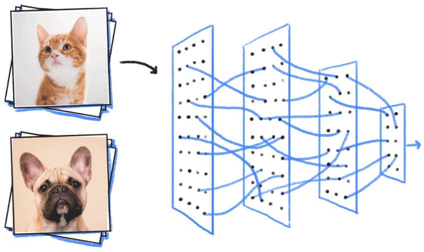
First of all, let's clear up the difference between equilibrium and nonequilibrium systems. If we find ourselves in a closed empty room we might talk about the temperature of the room which we call 'room temperature'. And that would be one number for the whole room; we wouldn't find any pockets of hot and cold air and even if we brought in a hot cup of coffee, heat would flow out of the coffee and eventually it would cool down to room temperature and reach *equilibrium* with the surrounding air in the room. In contrast, now let's imagine a room with a hot radiator on the left wall but it's December in Edinburgh and we've left a window on the right wall open. Heat will flow from left to right and while the radiator is on and window open, no matter how long we wait, that flow won't go away like it did for the cup of coffee. This is an example of an inherently *nonequilibrium* system - a system with a consistent flow in one direction.

A simple and widely studied model for such a system was proposed in 1968 by Macdonald *et al.* as a model for protein synthesis called the *totally asymmetric simple exclusion process* (TASEP) - forgive the name, it just means it is a model for particles jumping in one direction that can't land on top of each other nor overtake each other.

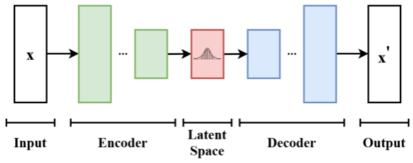


Now if we were talking about water, we might talk about its *phases* - two of which would be liquid water and ice. We could take our liquid water and place it in a freezer and when we come back later we would have ice. However, we did not just have all water one instant and all ice the next; we had a *coexistence* for a while - a bit of both. The TASEP also has phases, two of which are called low density (LD) and high density (HD). The LD phase means particles hop along unimpeded while in the HD phase, we see traffic jams. And we can induce our system to be in LD, HD or a permanent coexistence of the two. Intuitively, the physics of the coexistence is more 'complex'. We can simulate the TASEP and generate *snapshots* of the state of the system - our goal is to see how well an ML method can learn the *distribution* of these states ie. not focusing on any particular exact state but the likelihood of finding all the different states. If we were talking about an equilibrium system, the conversation ends here because any such system follows the same well-known distribution called the Boltzmann distribution.

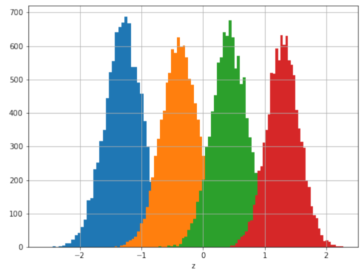
Our machine learning model of choice, the neural network (NN), are capable of anything from telling pictures of cats apart from dogs



to predicting the weather but it solves all these problems in the same way - learning features from data. NNs come in all shapes and sizes so our *architecture* of choice to learn the TASEP distribution is called the *variational autoencoder* (VAE). The algorithm here is not classifying or predicting anything, instead one half of the NN summarises our input data by *encoding* it by a small number of *latent variables* ie. what it finds 'important' about the data. The other half makes sure we are not overcompressing it so much that we loose the information and it does this by making sure it can still reconstruct the original data summarised by the latent variables.



Training VAEs on the TASEP snapshots and probing the latent variables showed that that the NN did learn what the important physical variables were and understood that the coexistence states were more complex than states that were purely LD or HD. It also managed to group different points of different densities together



To understand whether the VAE can learn more subtle features of the TASEP is not totally clear yet and further work could involve different nonequilibrium models and different NN architectures.

Understanding nonequilibrium systems has implications across all of STEM from biology to economics and although there is further work to be done, it has been shown that NNs could be a valuable tool to understand such systems.