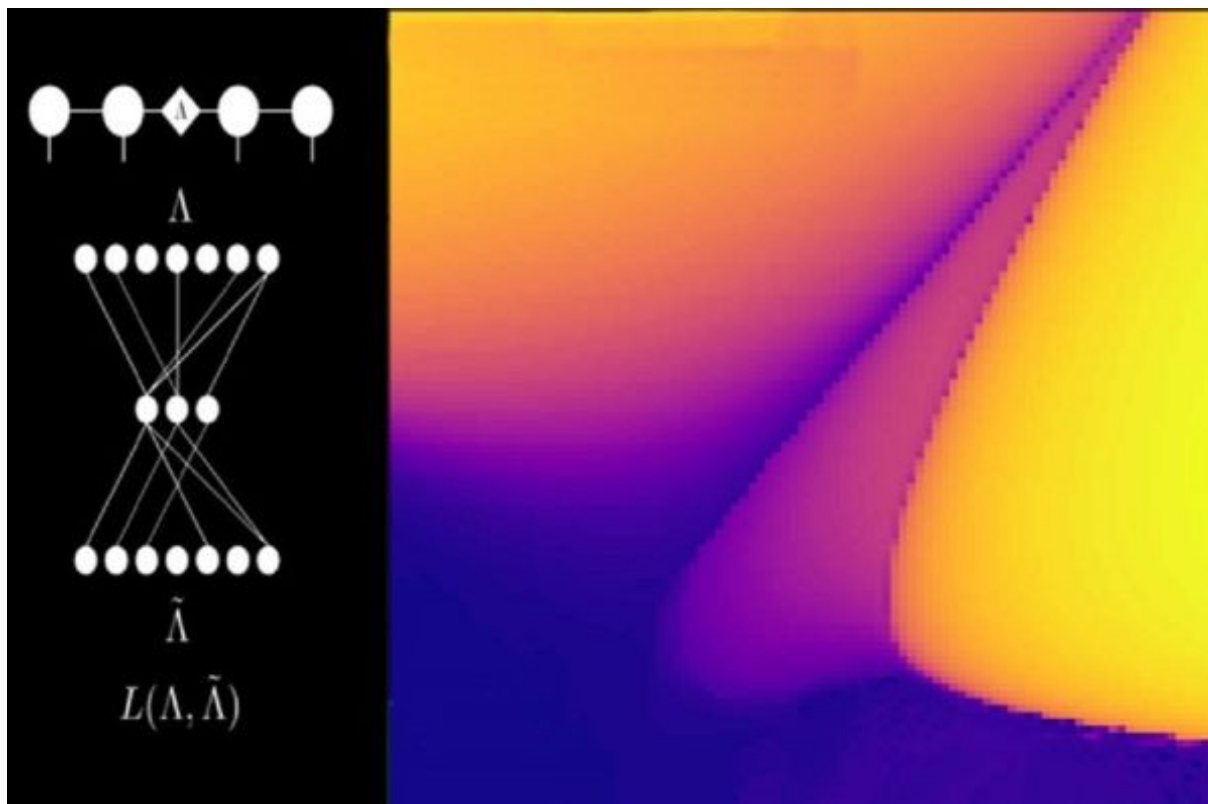




Novel approach of using Unsupervised Machine Learning in Physics

A team of ICFO researchers reports in PRL an entirely new anomaly detection method capable of training a system in very few iterations.

November 20, 2020

Machine Learning (ML) has the main goal of analyzing and interpreting data structures and patterns in order to learn from them, reason and carry out a decision-making task that is completely independent from human reasoning and engagement. Even though this field of study started in the mid 1900s, recent developments in the area have revolutionized the way on how we can process and find correlations in complex data.

Contrary to supervised learning, unsupervised learning seeks to discover patterns or classify information in large data sets into categories without prior knowledge. That is, it does not have labeled outputs, which means that it basically infers the natural structure that a dataset may have and extracts categorized information from it. This learning has proved to be very efficient for identifying phases and phase transitions of many-body systems.

In a study recently published in Physical Review Letters, ICFO researchers Korbinian Kottmann and Patrick Huembeli, led by ICREA Professors at ICFO Antonio Acín and Maciej Lewenstein report on a method that uses an unsupervised machine learning technique based on anomaly detection to automatically map out the phase diagram of a quantum many body system given unlabeled data.

The following example is very illustrative of what they have achieved. In machine learning the most common and known classification task example is to discriminate, for instance, images of cats and dogs. In the study, the method the researchers use anomaly detection, which handles the classification task of discriminating dogs and everything that is not a dog, approaching the system is an entirely different perspective. The idea is to train a special neural network called an autoencoder to efficiently compress and reproduce images of dogs. If the network is later fed with images of cats, the network does not know how to efficiently compress the features of the cat image and it is possible to tell from the higher reconstruction loss that it is not a dog.

This method is used in the context of quantum many-body systems. The images become observables, wave-functions or entanglement properties and the classes dogs and cats become different quantum phases. The model that they look at is the extended Bose-Hubbard model, which offers four different phases in the parameter space of interest. Since the researchers do not know the phases in their task a priori, they start by defining a region around the origin of the phase diagram as their starting point to train. Already from there, they are clearly capable of mapping the system in one training iteration, where all four phases of the system are easily distinguishable.

Finally, using this approach they have been able to reveal a phase-separated state between a supersolid and a superfluid, which appears in the system in addition to the standard superfluid, Mott insulator, Haldane-insulating, and density wave phases. The discovery of this new feature in this phase diagram is rather surprising provided that this model has been the investigated thoroughly for the past 20 years.

It is one of the first times in which a machine detects a previously unknown phenomenon in a quantum many-body context.