Abstract

The empirical success of neural networks at modeling highly complex relationships is supported by universality theorems which state that neural networks of the right size can approximate any function to arbitrary accuracy. These theorems pertain to the real numbers; in this thesis, we present a proof that neural networks with input, output, and parameters in a finite field can represent exactly any function over that field. Results are given for all finite fields except of orders two and four. In the process, we find activation functions which are suitable for such universality results to hold. However, optimization methods in this setting are not as well-defined as over the reals, which currently prevents us from being able to implement neural networks over finite fields in practice. Considering neural networks in this context highlights some of the challenges of machine learning over finite fields, and how it is fundamentally different from over real numbers. Machine learning algorithms over finite fields could have interesting applications in cryptography, coding theory, and computer science.