머신러닝을 이용한 연속함수와 이산함수에 대한 근사 가능성 분석

임형신[2], 권수진[2], 강주성[1,2], 염용진[1,2]*
국민대학교 정보보안암호수학과[1] / 금융정보보안학과[2]
{kuunh2, tnwls1595, jskang, *salt}@kookmin.ac.kr

An Analysis of Possibility of Approximations for Continuous and Discrete Functions Using Machine Learning

Hyoungshin Yim[2], Sujin Kwon[2], Ju-Sung Kang[1,2], Yongjin Yeom[1,2]*
Dept. of Information Security, Cryptology, and Mathematics[1] /
Financial information security[2], Kookmin Univ.

요 약

신경망을 이용한 연속함수의 근사 가능성은 Stone-Weierstrass 정리에 기반하여 다층 피드포워드 신경망이 Universal approximator된다는 것이 증명되었다. 하지만 이산함수의 근사 가능성은 이 증명에 의해 보장된다고 할 수 없다. 본 논문에서는 먼저 신경망을 이용한 연속함수의 근사 가능성을 보장하는 Universal approximator를 실험적으로 확인한다. 다음으로 이산함수의 경우에는 연속함수와 다르게 신경망을 이용한 근사 가능성이 성립하지 않을 수 있음을 구체적인 예를 통하여 입증한다.

Ⅰ. 서 론

신경망은 최근 컴퓨터 과학, 의학, 음성인식 등 여러 분야에서 많이 이용되고 있으며 많은 사람이 관심을 가지는 주제이다. 신경망은 단층신경망과 다층신경망으로 구분되며, 다층신경망은 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성된다. 신경망에 관한 많은 연구는 순방향 신경망(feedforward neural network) 구조에 따른 함수 근사 능력에 관한 세부주제로 초점을 맞추고 있다. 신경망을 이용한 연속함수의 근사 가능성은 신경망을 이용한 연속함수의 근사 가능성은 Stone-Weierstrass 정리에 기반하여 Hornik.K에 의해 다층 피드포워드 신경망(multilayer feedforward neural network)은 Universal approximator가 된다는 것이 증명되었다[1]. 이 연구는 하나의 은닉층을 가지는 충분히 큰 신경망은 원하는 정확도로 주어진 연속함수에 근사시킬 수 있음을 의미한다[1]. 한편, Stone-Weierstrass 정리는 이산함수에 대한 근사 가능성을 다루고 있지는 않다. 하지만 이산함수 관련 최근 신경망 연구 결과 중 일부는 이산함수의 근사 가능성은 논외로 한 상태에서 단순히 연속함수의 Universal approximator에 근거한 실험을 발표하였다[2]. 또한, 대표적인 이산함수인 부울함수(Boolean function)를 계산하는 데 필요로 하는 논리회로(logic gate)의 최소 개수를 제시한 아래의 Shannon 정리를 언급하지 않고 있다.

Shannon 정리[3]

There exists a Boolean function $g: \{0,1\}^n \rightarrow \{0,1\}$, on n variables, such that any circuit to compute f requires at least $\Omega(2^n/n)$ logic gates.

본 논문에서는 먼저 신경망을 이용한 연속함수의 근사 가능성을 실험적으로 분석하여 신경망이 Universal approximator임을 검증한다. 다음으로 이산함수의 경우에는 연속함수와 다르게 신경망을 이용한 근사 가능성이 성립하지 않을 수 있음을 구체적인 예를 통하여 입증한다. Shannon 정리에 나타난 대표적 이산함수인 부울함수는 신경망으로 근사시킬 때 필요로 하는 논리회로의 최소 개수가 필요하다. 이에 근거하여 부울함수의 신경망 구현에 필요한 논리회로의 개수를 산정하여, 이산함수는 연속함수의 경우와 달리 신경망을 이용한 Universal approximator와 같은 근사 가능성이 보장될 수 없음을 실험을 통하여 확인한다.

Ⅱ. 연속함수 및 이산함수의 신경망 근사 실험

본 절에서는 신경망 구조에 따른 연속함수 및 이산함수의 근사 가능성을 확인한다. 실험은 Windows 10 운영체제, Python 3.7.5 버전, Keras 2.3.1 버전에서 진행한다.

2.1. 실험 환경

표 1. 실험 데이터 수집 및 하이퍼 파라미터 설정

학습 데이터	연속함수 입력값	$[-1,1]$ 범위의 실수값을 랜덤하게 150개 추출		
	이산함수 입력값	0 또는 1을 랜덤하게 150개 추출		
테스트 데이터	연속함수 입력값	$[-1,1]$ 범위의 실수값을 랜덤하게 50개 추출		
	이산함수 입력값	0 또는 1을 랜덤하게 50개 추출		
에폭 수	1,000			
실험 차수	1차	2차	3차	
노드 개수	10^2	10^3	10^4	

표 1은 신경망의 연속함수 및 이산함수에 대한 근사 가능성을 실험하는 데 필요한 학습과 테스트 데이터 수집, 신경망의 하이퍼 파라미터(hyper parameter)를 정리한 것이다. 실험은 완전 연결 신경망(fully connected neural networks)을 사용한다. 높은 근사 정확도를 가지는 최소 에폭의 수가 1,000임을 실험적으로 확인하여 에폭의 수를 1,000으로 설정하였다. 다항식인 $f(x)=(x-0.3)(x+0.5)(x+1)(x-1)$을 근사시키고자 하는 연속함수로 설정한다.

이산함수는 $g:\{0,1\}^8 \rightarrow \{0,1\}$ 인 8-비트 입력에 1-비트 출력값을 가지는 비선형 함수로 $g(X) = A \bullet X^{-1} \oplus 1$ 로 표현되며, 추가적인 설명은 아래와 같이 서술한다.

- $X = a_7x^7 + a_6x^6 + a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x^1 + a_0x^0$는 이진 벡터 $(a_7, a_6, a_5, a_4, a_3, a_2, a_1, a_0)$ 로 표현되며, i가 $0 \le i \le 7$일 때, $a_i \in \{0,1\}$이다. X^{-1}는 $m(x) = x^8 + x^4 + x^3 + x + 1$ 상에서 X에 대한 역원을 의미한다.
- A는 $(1,1,1,0,0,0,0,1)$이며, $\bullet$은 내적(inner product)을 의미한다.

신경망에 의한 연속함수와 이산함수의 근사 가능성을 공정하게 비교해 보기 위하여 실험에 필요한 파라미터를 동일하게 설정한다. 이산함수는 Shannon 정리에 근거하여 적어도 $\Omega(2^8/8)$ 이상의 논리회로를 사용해야 하므로, 100개 이상의 노드 수를 사용한다. 연속함수에 관한 Universal approximator는 노드 수에 대한 대조군이 정의되지 않아 가변적으로 실험을 진행한다.

2.2. 실험 결과

본 실험은 연속함수와 이산함수의 근사실험 결과를 공정하게 비교하기 위하여 동일하게 L^2노름 값의 제곱으로 오차를 계산한다. L^2노름 값의 제곱에 대한 식은 다음과 같다.

$$(L^2 norm)^2 = \sum_{i=1}^{n} (y_i - \hat{y_i})^2.$$

이때, n은 테스트 데이터의 개수를 의미하며 y_i는 정답 레이블, $\hat{y_i}$는 예측된 값을 의미한다. 이산함수에 대한 오차는 전체 비트에서 잘못 예측한 비트의 비율을 계산하기 위해 0 또는 1로 반올림을 해준 뒤, 오차를 계산한다.

1) 연속함수에 대한 신경망 근사 실험

표 2. 연속함수 근사 실험에 따른 오차

실험 차수	1차	2차	3차
오차값 $(\times 10^{-4})$	2.49	1.09	8.10

학습 결과 구축된 신경망 함수의 출력값과 연속함수인 $f(x)$ 결괏값의 차이에 대한 L^2노름 값을 오차로 정의하여 얻은 결과가 표 2에 나타나 있다. 1차와 2차 실험의 결과를 통해 노드의 수가 증가할수록 오차가 감소하여 신경망으로 구축된 함수가 주어진 연속함수에 높은 정확도로 근사함을 확인하였다. 한편, 3차 실험에서는 단순한 함수에 비하여 상대적으로 너무 많은 노드가 사용됨으로써 과적합(overfitting) 현상이 발생한 경우이므로 신경망에 의한 근사는 적절한 노드 수의 설정이 필요하다는 사실을 발견할 수 있다.

2) 이산함수에 대한 신경망 근사 실험

표 3. 이산함수 근사 실험에 따른 오차율

실험 차수	1차	2차	3차
오차율(%)	62%	42%	58%

부울함수 $g(x)$ 값과 신경망으로 근사시킨 함수의 출력 비트값이 일치하지 않는 비율을 오차율로 정의하여 나타낸 결과가 표 3에 나타나 있다. 실험 결과에 나타난 오차율은 50% 근방에 흩어져 있음을 알 수 있다. 이는 0과 1을 무작위로 추출하는 랜덤한 부울함수의 오차율과 유사하여, 주어진 부울함수인 $g(x)$ 를 신경망으로 얻은 함수가 전혀 근사하고 있지 않음을 알 수 있다. 또한, 에폭 수를 변경하여 실험을 진행하여도 정확도가 개선되지 않음을 확인하였다.

Ⅲ. 신경망에 의한 이산함수의 근사 가능성

연속함수 및 이산함수의 실험 결과를 기반으로 신경망에 의한 이산함수의 근사 가능성에 관하여 고찰해보자. 표 2와 표 3의 2차 실험에서 에폭별 연속함수와 이산함수의 오차를 L^2노름으로 통합하여 계산한 결과는 그림 1에 나타나 있다. 학습이 반복될수록 연속함수의 오차는 0에 수렴한다. 한편, 이산함수의 손실은 0에 수렴하지 않고 25 근처로 계산된다. 부울함수에 대한 50개의 입력에 대한 함숫값 중에서 절반 정도가 다르게 나옴을 의미하여 신경망으로 근사시킨 함수가 랜덤함수와 별반 다르지 않음을 알 수 있다. 그러므로 이산함수의 신경망에 의한 근사는 연속함수와는 다른 방식으로 접근해야 할 것으로 보인다.

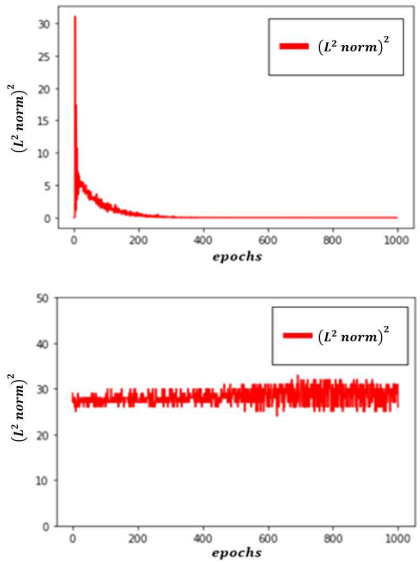

그림 1. 실험 2의 연속함수와 이산함수의 검증 손실

Ⅳ. 결론

본 논문에서는 하나의 은닉층을 갖는 신경망을 이용한 연속함수와 이산함수에 대한 근사 가능성을 실험적으로 분석하였다. 연속함수를 근사시킨 결과는 높은 정확도로 근사하였기 때문에 신경망이 Universal approximator임을 실험적으로 확인하였다. 한편, Shannon의 정리를 고려하여 연속함수와 동일한 조건으로 이산함수를 근사시킨 결과는 낮은 정확도를 보였다. 따라서 신경망을 이용한 이산함수의 근사는 연속함수와는 다른 방식으로 접근해야 할 것으로 보인다. 향후 연구로 신경망을 이용하여 이산함수인 암호의 평문을 복구하는 연구 결과를 검토하고, 신경망을 이용한 이산함수의 근사 방법을 주제로 연구할 예정이다.

참 고 문 헌

[1] Hornik K., Stinchcombe M., and White H. "Multilayer feedforward networks are universal approximators," Neural networks 2.5, pp. 359-366.

[2] Alani M. M. " Neuro-cryptanalysis of DES," World Congress on Internet Security (WorldCIS-2012). pp. 23-27. June, 2012

[3] J. Forbes, Nash, Jr., and M. Th. Rassias, "Open Problems in Mathematics." Springer, 2016