



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

# CNN 기반 원 달러 환율예측 모형 연구

A Study on CNN based Won-Dollar  
Exchange Rate Prediction Model



2016년

서강대학교 정보통신대학원

소프트웨어공학 전공

강 민 영

# CNN 기반 원 달러 환율예측 모형 연구

A Study on CNN based Won-Dollar  
Exchange Rate Prediction Model

지도교수 양 지 훈

이 논문을 공학석사 학위논문으로 제출함

2017년 6월 14일

서강대학교 정보통신대학원

소프트웨어공학 전공


강 민 영





## 논문인준서

강민영 의 공학석사 학위논문을 인준함.

2017년 6월 14일

주심 오경환 

부심 양외호 

부심 김지환 



# 목 차

제1장 서론 .....	1
1.1 연구의 배경 및 목적 .....	1
1.2 연구의 방법 및 구성 .....	2
제2장 이론적 배경 .....	3
2.1 환율 결정 이론 .....	3
2.2 환율 예측모형에 관한 이론 .....	6
2.2.1 ARIMA 모형 .....	6
2.2.2 인공신경망 모형 .....	7
2.3 환율예측에 관한 기존 연구 검토 .....	12
2.4 합성곱 신경망(Convolutional Neural Network) 모형 .....	14
제3장 합성곱 신경망(CNN)을 이용한 모형 구축 .....	20
3.1 연구의 대상 .....	20
3.2 자료수집 .....	21
3.3 실험데이터 전처리 .....	24



제4장 연구 결과 및 해석 .....	30
4.1 선행 연구 모형 평가 .....	30
4.1.1 Multi Layer Perceptron 모델 .....	30
4.1.2 Multi Layer Perceptron 실험 결과 .....	31
4.2 구축 모형에 대한 예측과 평가 .....	32
4.2.1 Convolutional Neural Network 모델 .....	32
4.2.2 Convolutional Neural Network 실험 결과 .....	34
4.3 실험 결과 분석 .....	35
제5장 결론 및 향후 과제 .....	36
5.1 결론 .....	36
5.2 향후 과제 .....	38
참고 문헌 .....	39



## 표 목 차

[표 2-1] 원화 환율에 영향을 미치는 다양한 요인들 .....	4
[표 3-1] 분석에 사용된 지표 .....	21
[표 3-2] 원본 데이터의 예 .....	23
[표 3-3] 데이터 분할 기간 .....	24
[표 3-4] 선형 보간법 적용 후 데이터 .....	27
[표 3-5] 정규화 후 데이터의 예 .....	28
[표 3-6] Input Data .....	29
[표 4-1] MLP 모델 실험 결과 .....	32
[표 4-2] Convolution 2-Layer 적용시 매개변수 연구 결과 .....	32
[표 4-3] Convolution 1-Layer 적용시 매개변수 연구 결과 .....	34
[표 4-4] 성능 측정 방법별 성능 .....	35



## 그 립 목 차

[그림 2-1] 단층신경망의 구조 .....	8
[그림 2-2] 다층신경망의 구조 .....	9
[그림 2-3] Feed Forwad 계산 방법 .....	10
[그림 2-4] Sigmoid함수와 Relu함수 .....	11
[그림 2-5] Back Propagation 동작 방식 .....	12
[그림 2-6] LeNet-5(CNN)의 구조 .....	15
[그림 2-7] FC 와 CNN의 구조적 차이 .....	15
[그림 2-8] 합성곱(CNN) 연산 예제 .....	16
[그림 2-9] 출력 크기 계산식 .....	17
[그림 2-10] Max-pooling 예제 .....	18
[그림 3-1] 선형보간법 그래프 .....	25
[그림 3-2] 선형보간법 그래프(일반화) .....	25
[그림 3-3] 선형보간법 수식 .....	26
[그림 3-4] 선형보간법 수식(정규화) .....	26





[그림 4-1] MLP 모델의 구조 .....	30
[그림 4-2] MLP 모델 MATLAB 실험 과정 .....	31
[그림 4-3] Convolution 2-Layer 적용시 CNN 최적 모델 구조 .....	33
[그림 4-4] Convolution 1-Layer 적용시 CNN 최적 모델 구조 .....	34



# Abstract

High-tech industry requires a high level of technology and high-skilled manpower. The manufacturing environment is global and the manufacturing cost is very high. It requires a large-scale investment in technology and production facility implementation. Therefore, the importance of accurate price forecasting cannot be emphasized enough as its failure can incur a huge loss due to the nature of the industry. High-tech industry has a large proportion of exports and imports, and changes in raw material prices due to exchange rates can result in fluctuations in production cost. Therefore, a huge number of previous studies have constructed exchange rate forecasting models. However, since existing statistical or econometric models showed certain limitations in the analysis of nonlinear time-series data such as stock prices and exchange rates, an artificial neural network has been proposed as an alternative analytical model of time-series data. Moreover, with the recent development of various types of deep learning techniques to supplement the problems of artificial neural networks, this study aims to search for a new model through comparison with previous research models and to predict and evaluate the construction model.

The paper predicted the exchange rate using Deep Learning



which showed superior performance in various fields such as image recognition, speech recognition, and natural language processing. This paper used the convolutional neural network (CNN) to solve the various shortcomings of the deep neural network among the deep learning algorithms and attempted to predict the exchange rate by selecting six variables that could affect the exchange rate. The data used for the analysis is daily data from January 3, 2000 to December 31, 2016. In order to construct the model, it was divided into two parts: (1) training data and (2) verification data to verify the constructed model. The prediction of the exchange rate by inputting the verification data into the learned neural network model showed a certain extent of difference to the actual value. However, it was possible to figure out the trend of the exchange rate and it was better than or similar to the previous studies modeled by the artificial neural network.



## 초 록

고도의 기술력과 인력을 요구하는 첨단 산업은 제조 환경이 글로벌화 되어있을 뿐만 아니라 제조 단가가 매우 높다. 또한 대규모의 기술 투자 및 대규모 생산시설을 구축하기 위해 매우 큰 투자가 요구되어 산업 특성상 실패 했을 경우 큰 손실을 입게 되므로 정확한 가격 예측이 중요하다 할 수 있다. 수출과 수입의 비중이 높은 산업으로 환율에 따른 원자재 가격의 변화는 생산 단가의 변동으로 이어지게 된다. 따라서 그동안 수많은 선행 연구에서 환율 예측 모형을 구축하였으나 기존의 통계적 또는 계량경제 모형이 주식가격, 환율 등과 같이 예측하기 어려운 비선형적인 특성을 가지는 시계열 자료의 분석에 일정한 한계를 보임에 따라 시계열 자료의 대안적 분석 모형으로 인공신경망이 제시되었다. 더 나아가 최근에는 인공신경망의 문제점을 보완하기 위해 딥러닝 기법이 다양하게 발달하고 있는바 이에 본 논문은 최근 이미지 인식, 음성 인식, 자연어 처리 등 다양한 분야에서 우수한 성능을 보여주고 있는 합성곱 신경망(Convolutional Neural Network)을 환율 예측에 적용하여 시계열 데이터 분석에서의 성능을 평가하는 것을 목적으로 한다. 딥러닝 알고리즘 중 신경망의 여러 단점들을 해결해주기 위해 합성곱 신경망을 이용하였으며 환율에 영향을 줄 수 있는 변수 6개를 선택하여 환율의 예측을 시도하였다. 분석에 사용한 데이터는 2000년 1월 3일부터 2016년 12월 31일까지의 일별 데이터이다. 모형 구축을 위하여 학습용 데이터와 구축된 모형의 검증을 위한 검증용 데이터의



두 부분으로 분할하였다. 검증용 데이터를 학습된 신경망 모형에 입력하여 환율을 예측한 결과, 실제 값과 어느 정도의 차이를 보여주었지만 환율의 흐름을 파악하는 것이 가능하였고 인공신경망으로 모델링한 선행 연구와 비교했을 때 좋은 성능을 보였다.



# 제1장 서론

## 1.1 연구의 배경 및 목적

통신기기, 정보기기, 자동차, 군사, 항공 등 첨단 산업은 수출, 수입의 비중이 높은 산업인 만큼 환율에 따라 기업의 이익이 크게 차이가 날 수 있기 때문에 기업에서는 환율을 예측하여 기업의 이익안정화를 위해 노력하고 있다. 환율 예측은 기업에서 투자를 할 때, 매우 중요한 척도로서 기술 및 생산에 투자하였을 경우, 장기적인 손실과 이익을 예측할 수 있도록 해준다. 따라서 기업에서는 환율 예측이 무엇보다 중요해졌고 전문가의 논리적 판단과 경험을 기반으로 하는 시스템 보다는 정확한 환율 결정을 지원하는 시스템을 필요로 하며 환율에 영향을 주는 주요 인자 및 환율 예측 모형 연구를 필요로 하게 되었다.

뒤늦게 많은 사람들이 환율 예측에 대해 연구를 하고 많은 모형들을 내놓았지만 기존의 통계적 또는 계량경제 모형으로는 환율 예측을 하는데 많은 한계점을 가지고 있었다. 따라서 이러한 한계점을 해결하고자 비선형 시계열 데이터의 분석에 우수한 신경망 이론들이 나오게 되었고 활발히 연구가 되고 있다. 본 논문은 신경망이론 중 합성곱 신경망(Convolutional Neural Network) 알고리즘을 통해 영상과 음성 외에 시계열 데이터 등에도 좋은 성능을 보일 수 있는지 연구해 보고자 해당 알고리즘을 적용하여 모형을 구축하고 기존 환율 예측 모형과 비교해 보아 해당 모형을 평가하고자 한다.



## 1.2 연구의 방법 및 구성

본 논문에서는 환율 예측시 복잡하고 일반화된 공식을 사용하지 않고 신경망 알고리즘을 이용하여 해결하고자 시도하였다. 원-달러 환율 예측을 신경망 알고리즘을 적용하였고 환율 결정에 영향을 미치는 요소를 입력변수로 선택하였다. 기존 연구에는 주로 금리, 주가 등의 시계열 요소가 입력변수로 선택되었으나 거시 경제변수가 환율 결정에 미치는 영향을 알아보기 위해서 자본수지, 경상수지, 경제성장률, 물가도 추가 되었다. 실증적 자료를 통한 분석은 과거 15년간의 데이터를 이용하여 일별 원-달러 환율 예측을 시도하였다.

본 논문은 총 5장으로 구성되어 있다. 1장에서는 연구의 배경 및 목적에 대하여 서술하고 2장에서는 환율 예측모형에 관한 이론을 서술하고 인공신경망과 딥러닝에 대한 이론을 살펴보고 선행연구에 대해 고찰한다. 3장에서는 변수 선택 및 연구 모형 구축에 대해 서술하고 4장에서는 실제 자료 분석 결과를 보여준다. 마지막 5장에서는 연구의 결론과 향후 과제를 서술한다.



## 제2장 이론적 배경

### 2.1 환율 결정 이론

환율이란 자국화폐와 타국화폐간의 교환비율을 말하는 것으로, 환율은 일반적으로 수요와 공급의 법칙에 따라 결정된다.

외화에 대한 수요가 공급을 초과하면 외화의 가치는 상승하여 원화 환율은 상승하고, 공급이 수요를 초과하면 외화의 가치는 하락하여 원화 환율은 하락한다. 우리나라 외환시장에서도 외화의 수요가 늘어나는 경우는 해외에서 상품을 수입해서 외화로 결제하는 경우, 여행경비나 차관이자를 외화로 지급하는 경우가 있다. 반대로 외화의 공급이 늘어나는 경우는 상품을 수출하고 대금을 외화로 받을 경우, 외국에 나가 일해서 그 대가를 외화로 받을 경우, 외국인 관광객이 한국에 와서 외화를 사용할 경우, 외국인이 국내에서 주식투자를 하기위해 외화를 원화로 바꾸는 경우 등이 있다. 이러한 일련의 경제활동 결과를 종합하여 원화 수요가 공급을 초과한다면 원화 환율은 상승하고, 그 반대의 경우는 원화 환율이 하락한다[3].

[표 2-1]은 원화 환율에 영향을 미치는 다양한 요인들에 대해 정리한 것으로 이 중 일부는 본 논문의 연구 대상의 입력변수가 될 것이다.





[표 2-1] 원화 환율에 영향을 미치는 다양한 요인들[3]

물가	<p>① 장기적인 환율의 결정요인으로 자국의 물가상승률이 타국에 비하여 상대적으로 상승 추세에 있을 때 국내 상품에 대한 수요가 하락하고 외국상품에 대한 수요가 상승하여 자국의 수출경쟁력은 약화되는 반면, 수입수요는 증대되어 경상수지의 악화를 유발시키며 이로 인하여 자국 통화의 가치 하락을 만들게 됨.</p> <p>② 소비자들은 물가인상이 자신의 구매력에 끼치는 영향력을 피하고자 높은 물가인상률을 보이는 국가로부터 수입되는 상품보다 낮은 물가인상률을 보이는 국가로부터 수입되는 상품에 보다 큰 관심을 보인다. 따라서 낮은 물가인상률을 보이는 국가의 통화가치는 상승하는 반면, 높은 물가인상률을 보이는 국가의 통화가치는 하락함.</p>
정치적 요인	정치적, 경제적으로 불안한 경우 해당국의 통화를 기피하고 다른 안정된 통화로 투자대상을 바꾸려는 투자자들 때문에 해당통화는 약화되며, 안정된 통화를 선호하는 경향에 따라 정치적, 경제적으로 안전한 국가가 강세를 나타내기도 함.
금리 수준	각국의 상대적 금리수준의 변동은 다른 조건이 일정할 때 환율변동의 주요 요인이 된다.
통화량	통화량의 증가는 다른 조건이 일정할 때 자국 금리의 하락요인이며 인플레이션의 상승요인이 되어 자국통화의 가치하락을 유발시킨다.



시장 분위기	<p>시장분위기는 때로 이성적이고 합리적인 의사결정 수단과 정보를 무력화시키는 경우가 있는데 경제사정에 따라 외환시장에서도 이러한 경향은 언제든지 나타날 수 있다. 일반적으로 외환시장 참가자들의 심리는 민감하며, 참가자들은 환율에 영향을 미치는 각종 경제지표 등을 미리 예측하여 거래를 하게 된다. 따라서 국제정치의 급변 및 석유 등의 에너지 문제 따위가 발생하면 그 움직임은 환율에 즉각적으로 반영된다. 또한 환율의 전망이 불투명한 상태에 있을 때 통화 당국자나 외환시장의 유력한 조언 리더의 발언이 환율을 움직이게 만들기도 한다. 실제 한국은행이나 경제관련 당국자의 외환시장 개입의지 피력정도에 따라 환율의 차이를 가져오기도 한다.</p>
당국의 외환시장 개입	<p>통화당국은 불안한 환율을 안정시킬 목적으로 외환시장에 개입하는데 이러한 외환시장 개입은 환율에 영향을 미치며, 정책을 구사함으로써 환율 변동으로부터 자국통화를 보호하기도 한다.</p>
환경요인	<p>전염병이나 태풍 등 재해의 발생</p>
기대수준	<p>각 경제주체들은 국제외환시장의 동향과 통화량, 국제수지 등 거시경제에 대한 전망, 그리고 당국의 정책방향 등 환율에 영향을 미칠 수 있는 여러 가지 요인들에 대한 정보를 수집하고 이를 토대로 장래 환율에 대한 기대를 형성하고 위치를 조정함.</p>



## 2.2 환율 예측 모형에 관한 이론

### 2.2.1 ARIMA 모형

시계열 모형을 이용한 환율예측모형은 환율변동의 역사적 추이에서 특정 모형을 추출하여 미래의 환율을 예측하고자 하는 모형이다. 즉, 구조적 모형을 이용하여 미래의 환율을 예측할 경우 독립변수들의 미래 예측치를 추정하거나 예상하는 어려움을 해결하기 위해 시계열분석은 특정변수의 미래값을 예측하기 위하여 단지 해당변수 자체의 과거 변동유형을 분석하여 미래의 모습을 가장 가깝게 재생하는 것이다. 이 중 대표적인 것이 ARIMA 모형이다.

시계열 분석방법 중 Box-Jenkins에 의해 개발된 ARIMA모형을 확률적 시계열 모형이라고 한다. 자기회귀(Autoregressive)와 이동평균(Moving Average) 등 대부분의 시계열이 불안정하기 때문에 이를 안정화시키기 위한 차분의 과정이 포함된 확률적 시계열 모형이 ARIMA모형이다.

ARIMA 모형을 이용한 환율예측에는 외생변수와 같은 개념이 존재하지 않기 때문에 미래의 예측기간에 대한 외생변수의 사전적인 결정이 필요하지 않다. 그러나 외생변수의 개념이 없으므로 다른 경제 변수들의 효과를 설명하는 것이 불가능하다는 단점이 있다[3].



### 2.2.2 인공신경망 모형

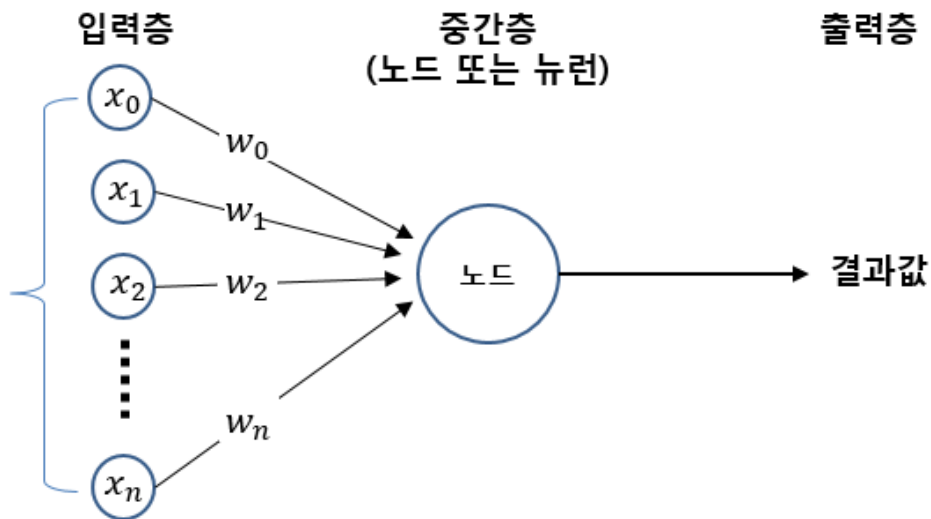
신경망은 인간의 두뇌의 신경 세포를 모방한 개념으로 마디(Node)와 고리(Link)로 구성된 망구조를 모형화하고, 과거에 수집된 데이터로부터 반복적인 학습과정을 거쳐 데이터에 내재되어 있는 패턴을 찾아내는 모형화 기법이다. 신경조직의 기본적인 구성요소는 신경세포 즉, 뉴런(Neuron)이다. 인공신경망은 인간두뇌의 정보처리과정과 그 구조를 모형화 한 것으로서 이는 뇌의 신경회로망같이 신경세포에 해당하는 다수의 처리 프로세스들을 네트워크 형태로 접속하여 상호간에 신호를 주고 받으면서 병렬로 작동하여 하나의 인공신경망으로 자료를 처리한다. 그 구조는 입력층(Input Layer), 은닉층(Hidden Layer), 출력층(Output Layer)의 삼층구조로 나뉘지며, 각 층마다 정보처리단위(Processing Unit)인 다수의 뉴런을 가지고 있다[4]. 은닉층이라고 불리는 독특한 구성요소에 의해서 일반적인 통계모형과 구별되어 진다. 은닉층은 인간의 신경세포를 모형화한 것으로 각 은닉층은 입력변수들의 결합을 수신하여 목표변수에 전달한다. 이때 결합에 사용되는 계수(Coefficient)들을 연결강도(Synaptic)라고 부르며, 활성화함수는 입력값을 변환하고 이를 입력으로 사용하는 다른 마디로 출력하게 된다.

인공신경망은 인간의 뇌 구조에서 영감을 얻은 학습 기법이다. 인간의 뇌에는 수백억 개의 신경 세포인 뉴런으로 구성되어 있으며, 뉴런은 시냅스들로 조밀하게 연결되어 있다. 특정 뉴런에 신호가 전달되면 그 뉴런은 시냅스를 통해 다른 뉴런으로 신호를 전달하고 이 과정을 통해 뇌는 새로운 것을 학습한다. 인공신경망은 뉴런을 노드



로 이용하고 시냅스를 가중치로 결합하여 네트워크를 구성한다.

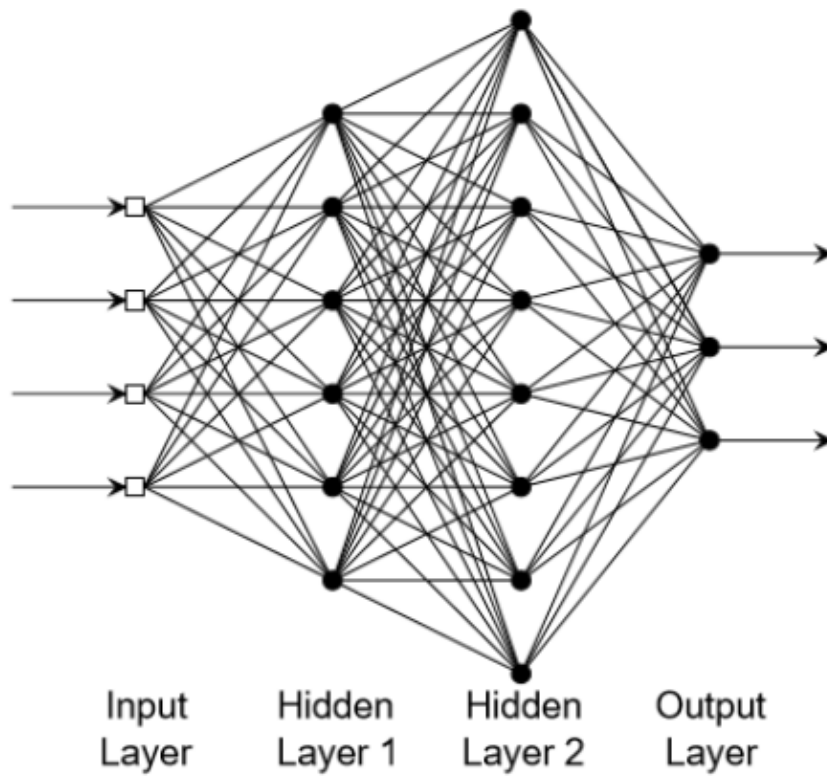
인공신경망은 Single Layer와 Multi Layer로 나뉜다. Single Layer은 가장 간단한 구조로서 [그림 2-1]과 같이 입력을 받아들이는 입력층, 중간층, 신경망의 결과가 출력되는 출력층으로 구성된다.



[그림 2-1] 단층신경망의 구조

Multi Neural Network는 [그림 2-2]와 같이 여러 층을 가진 신경망이다. 일반적으로 3계층 신경망을 주로 사용하며, 입력을 받는 입력층(Input), 신경망의 결과가 출력되는 출력층(Output), 입력층과 출력층 사이에 외부로 나타나지 않는 은닉층(Hidden)의 3계층으로 구성된다. 그리고 두 개 이상의 은닉층이 있을 경우에 Deep Neural Network라고 표현한다.

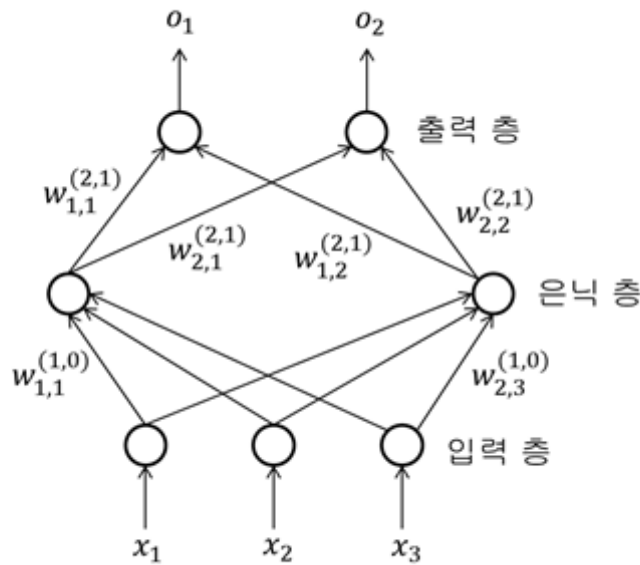




[그림 2-2] 다층신경망의 구조[4]

신경망을 훈련시키는 방식에 Feed Forward와 Back Propagation이 있다. 먼저 Feed Forward는 신경망의 네트워크가 Forward 방향으로 뉴런의 값을 전달하여 계산하는 방식이다.

각 뉴런은 [그림 2-3]과 같이 계산이 된다.



[그림 2-3] Feed Forward 계산 방법

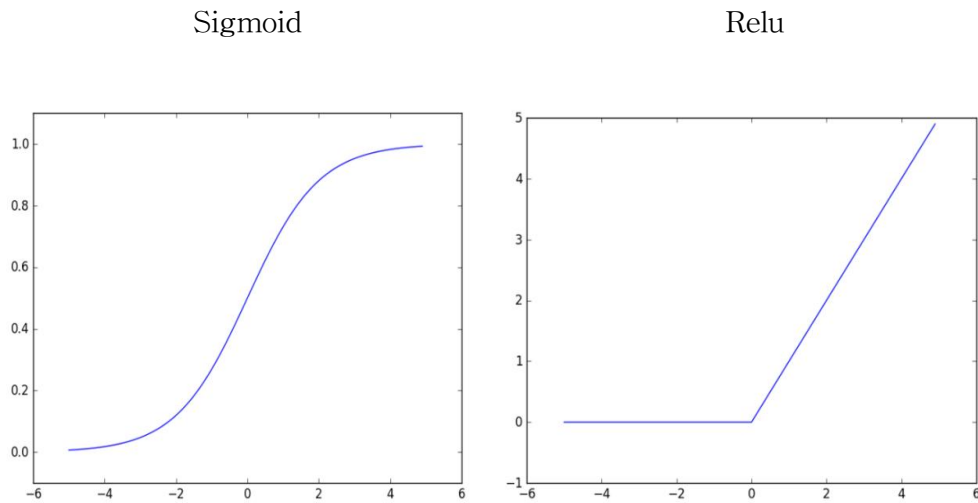
신경망에서는 노드와 가중치의 합은 활성화 함수활성화 함수를 통해 출력된다. 활성화 함수는 주로 Sigmoid함수와 ReLu함수가 주로 사용된다.

[그림 2-4]와 같이 Sigmoid 함수는 신경망의 출력을 곡선 형태로 바꾸어주며, 연속적이며 미분 가능하도록 만들어 준다. 또한 신경망의 출력값을 일정한 값으로 수렴시켜 준다. Sigmoid는 S자 모양이라는 뜻이다. 이렇게 S자 모양을 가지는 함수들의 집합을 Logistic Function이라고 하는데, Sigmoid Function은 Logistic Function에서  $L=1$ 이고 Midpoint가 0인 경우를 말한다.

최근에는 Sigmoid 함수대신 Relu함수를 많이 사용한다. 선형함수



이며, 미분항이 0 또는 1이 나와 계산량을 줄여준다. 신경망의 층이 많아질 때 속도가 느려진다는 단점을 해결해 준다.

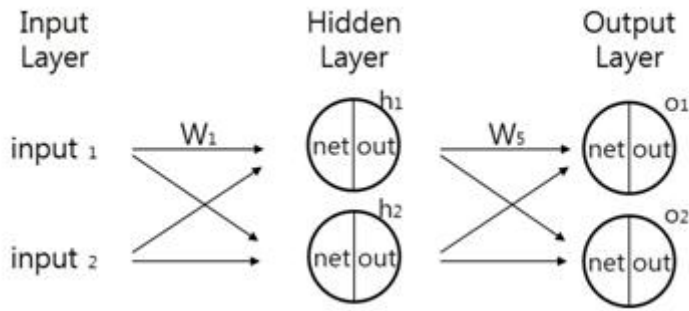


[그림 2-4] Sigmoid함수와 Relu함수[5]

또 다른 오류 Back Propagation (역전파 알고리즘)은 이름에서 알 수 있듯 출력층의 결과값에서부터 역으로 가중치를 학습 시키게 된다. Back Propagation 의 기본 아이디어는 우리가 은닉층 그 자체에 대해서 알 필요가 하나도 없고 대신 은닉층으로 인해 생성되는 오류의 변화를 관측하는 것이 더 낫다는 것이다. 즉, 은닉층 그 자체를 학습하는 것이 아니라, 은닉층으로 인해서 생겨나는 오류를 사용하는 것이다. [그림 2-5] Back Propagation은 주어진 가중치 조정을 위해 입력에 대한 출력과 목표 출력을 이용해서 오차E를 구한다.







[그림 2-5] Back Propagation 동작 방식

### 2.3 환율예측에 관한 기존 연구 검토

김재현[3]은 환율예측을 위한 방법으로 인공신경망 모형이나 ARIMA 모형을 이용한 기존의 환율예측연구에 이어 두 모형의 원/달러 환율 예측력을 비교하여 어느 모형이 우수한 예측성과를 보이는지를 연구하였다. 이를 위하여 시장평균환율제도인 1992년 7월 1일부터 1994년 10월 31일까지의 기간을 연구기간Ⅰ, 1994년 11월 1일부터 1996년 12월 31일까지의 기간을 연구기간Ⅱ, 그리고 변동환율제도인 1999년 1월 1일부터 2000년 12월 31일까지의 기간을 연구기간Ⅲ으로 연구기간을 구분하다. 그리고 일일 원/달러 환율을 사용하여 각 기간별로 인공신경망 모형과 ARIMA 모형을 구축하여 가장 적합한 환율예측모형을 선정하다. 그 후 각 연구기간별로 선정된 인공신경망 모형과 ARIMA 모형을 이용하여 마지막 90일의 환율을 예측하다. 마지막으로 예측결과에 대해 어느 모형이 보다 우수한 환율예측성과를 보이는지를 평균자승오차근과 평균절대오차의 비교, 환율변동방향



적중율의 비교, 그리고 Ashley, Granger and Schalensee(1980)의 두 모형간의 예측오차에 대한 회귀분석을 실시함으로써 실증 분석하였다.

두 모형의 환율변화율점 예측 정확성에 대한 비교에서는 평균자승 오차근과 평균절대오차를 기준으로 세 예측기간 모두 인공신경망 모형이 평균적으로 우수한 예측력을 보이는 것으로 나타났다. 또한 두 모형간의 평균자승오차근과 평균 절대오차의 차이가 자유변동환율제도인 예측기간Ⅲ에서 가장 큰 것으로 나타나 시장평균환율 제도에 비하여 자유변동환율제도에서 두 모형간의 평균적인 환율예측오차의 차이가 확대되었다. 따라서 일일 원/달러 환율예측에 있어서 인공신경망 모형이 ARIMA 모형보다 평균적으로 우수한 환율 예측력을 보인다고 할 수 있었다.

이경선[7]은 미래의 원-달러를 예측함에 있어서, 인공지능의 한 분야인 신경망을 이용하였다. 신경망은 인간의 뇌 구조를 이용하여 모델링한 알고리즘으로 반복된 입력정보에 대하여 목적에 맞도록 변화시킴으로써 훈련을 할 수 있다. 환율예측 문제를 신경망을 이용함으로써 비교적 높은 성과를 제시할 수 있도록 하는 것이 연구의 목적이다. 신경망의 입력변수는 환율결정에 유의한 변수들을 선택하였는데 특히 거시경제변수가 환율결정에 미치는 영향을 알아보기 위해 금리, 인플레이션, GDP성장률, 국제 수지와 주가지수 등의 일반 시계열 변수를 선택하였다. 다층 퍼셉트론을 이용한 역전파 알고리즘으로 모형을 훈련시키고 예측한 결과 실제환율과의 평균오차가 11.24원으로 비교적 안정된 결과를 보여주고 있었다. 거시적인 경제 요인들이 환율변화에 영향을 준다는 사실을 근거로 거시경제변수와 시계열 변



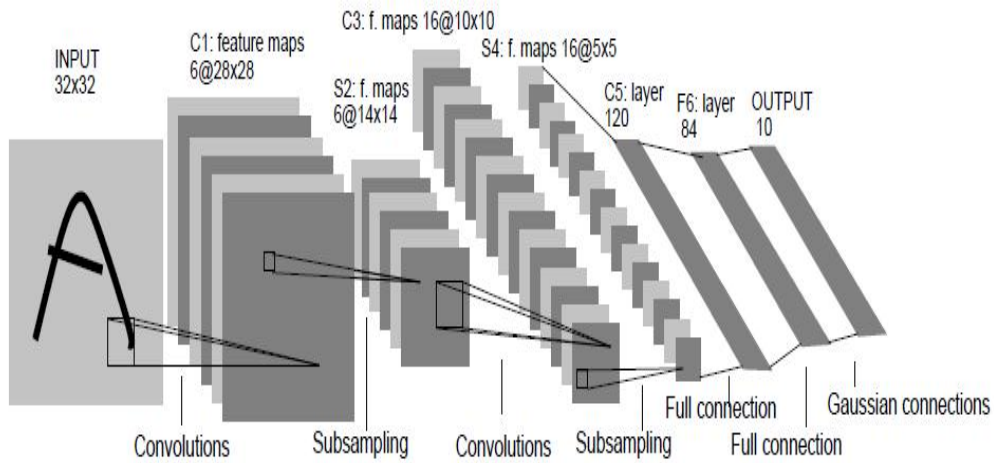
수를 분리하여 따로 분석하여 예측한 결과 거시경제변수로 구성된 모형이 일반 시계열 변수 모형보다 3배 정도 나은 예측력을 보여주고 있다. 다중회귀분석과 신경망에 의한 예측오차를 비교를 해도 거시경제변수를 이용한 신경망에서 도출한 예측력이 향상적임을 보이고 있다. 따라서 본 연구에서 제안한 신경망 모형은 비교적 안정된 예측력을 보임으로써 의사결정을 위한 활용도구로 가치를 가지고 있다고 볼 수 있었다.

## 2.4 합성곱 신경망(Convolutional Neural Network) 모형

합성곱 신경망(CNN, Convolutional Neural Network)는 인간이 사물을 인식할 때 시각적 정보를 바탕으로 특징을 추출한다는 점에 기반하여 개발된 알고리즘이다. CNN의 특징은 Sparse Connectivity, Shared Weight이다. 인접한 신호들에 대해서만 필터를 적용하여 기존의 입력데이터를 전부 처리할 필요가 없어 계산량을 감소시킨다. 또한, 필터를 Weight로 입력 데이터 전체에 반복적으로 적용하여 변수의 수를 줄인다. CNN은 필터를 적용하여 Feature Map을 추출하는 Convolution Layer, Feature의 수를 줄이는 Pooling Layer, 모든 노드들이 연결되어 있는 Fully Connected Layer로 구성된다.

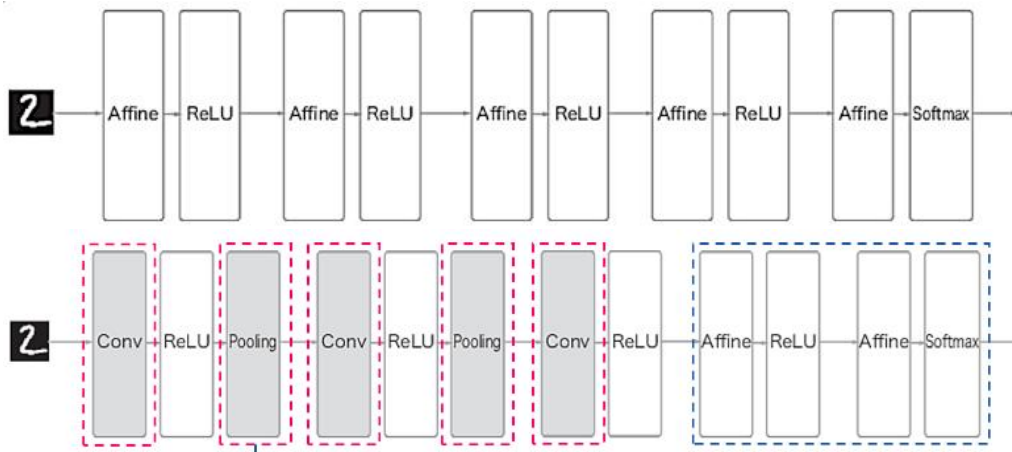
다음 [그림 2-6]는 Lecun이 발표한 논문에 실린 CNN의 구조도이다[5].





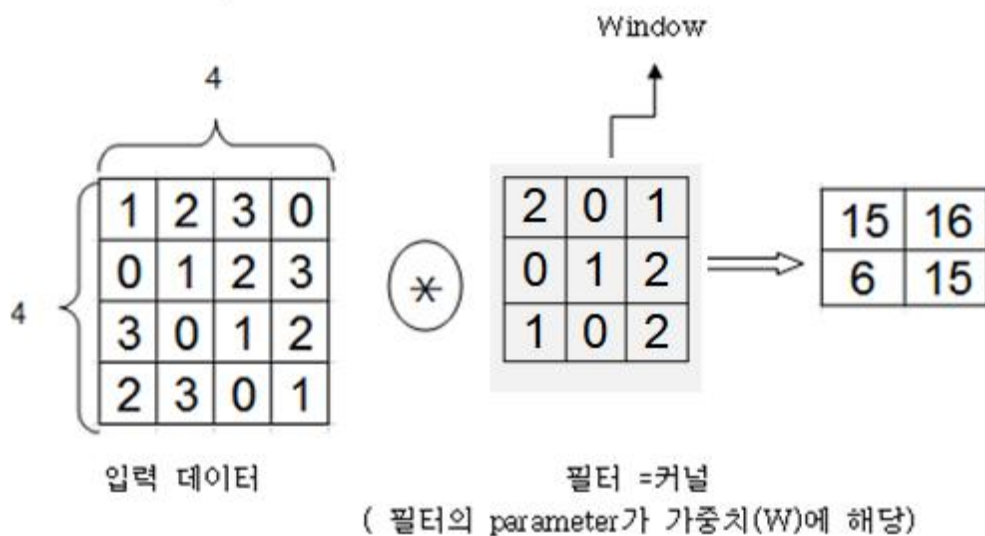
[그림 2-6] LeNet-5(CNN)의 구조[5]

CNN의 구조는 완전연결계층(Fully-Connected Layer)과는 다른점이 있다. Fully-Connected Layer에서는 Affine계층으로 구현했지만, CNN에서는 아래의 [그림2-7]과 같이 합성곱 계층(Convolutinal Layer) 과 풀링 계층(Pooling Layer)이 추가된다[6].



[그림 2-7] FC와 CNN의 구조적 차이[6]

합성곱 계층에서 연산은 데이터와 필터의 모양을 (높이, 너비)로 나타내고, 윈도우(Window)라고 부른다. [그림 2-8]를 통해 계산 방법을 알아보면 입력데이터는 (4, 4), 필터는 (3, 3)이고, 필터가 바로 Conv Layer의 가중치에 해당한다. 합성곱 연산은 필터의 윈도우를 일정한 간격으로 이동해가며 계산한다. 합성곱 연산은 입력데이터와 필터 간에 서로 대응하는 원소끼리 곱한 후 총합을 구하게 되며, 이것을 Fused Multiply-Add 라고 한다. 마지막으로 편향은 필터를 적용한 후에 더해지게 된다[6].



[그림 2-8] 합성곱 연산 예제[6]

패딩(Padding)은 합성곱 연산을 수행하기 전, 입력데이터 주변을 특정 값으로 채워 늘리는 것을 말한다. 패딩(Padding)은 주로 출력데이터의 공간적(Spatial)크기를 조절하기 위해 사용한다. 패딩을 할때, 채울 값은 매개변수로 어떤 값을 채울지 결정할 수 있다. 주로

Zero-Padding을 사용한다. 패딩을 사용하는 이유는 패딩을 사용하지 않을 경우, 데이터의 공간크기는 합성곱 계층(Conv Layer)를 지날 때 마다 작아지게 되므로, 가장자리의 정보들이 사라지는 문제가 발생하기 때문에 패딩을 사용한다.

스트라이드(Stride)는 입력데이터에 필터를 적용할 때 이동할 간격을 조절하는 것, 즉 필터가 이동할 간격을 말한다. 스트라이드 또한 출력 데이터의 크기를 조절하기 위해 사용한다. 스트라이드(Stride)는 보통 1과 같이 작은 값이 더 잘 작동하며, 스트라이드가 1일 경우 입력 데이터의 크기는 풀링계층에서만 조절하게 할 수 있다.

패딩과 스트라이드를 적용하고, 입력데이터와 필터의 크기가 주어졌을 때 출력데이터의 크기를 구하는 식은 아래 [그림 2-9]과 같다 [6].

$$(OH, OW) = \left( \frac{H + 2P - FH}{S} + 1, \frac{W + 2P - FW}{S} + 1 \right)$$

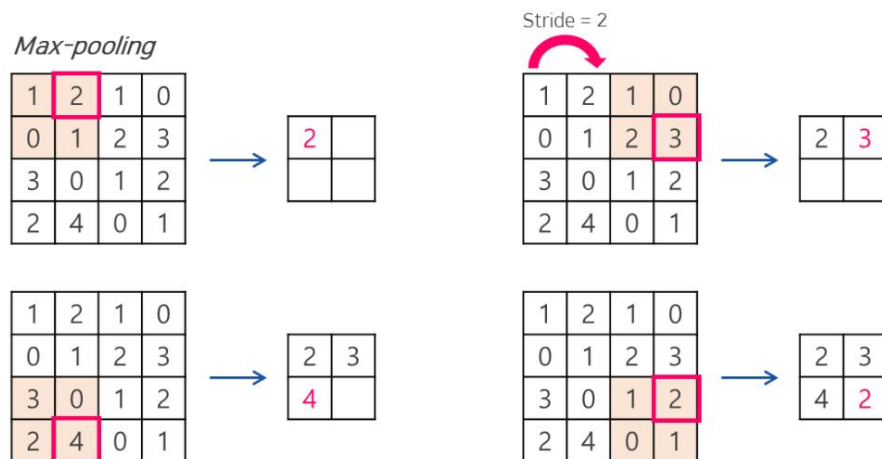
- $(H, W)$  : 입력크기
- $(FH, FW)$  : 필터크기
- $(OH, OW)$  : 출력크기
- $P$  : 패딩
- $S$  : 스트라이드

[그림 2-9] 출력 크기 계산식[6]

풀링계층은 합성곱 계층의 패딩과 스트라이드처럼 데이터의 공간적 크기를 축소하는데 사용한다. 주로 합성곱 계층(Conv Layer)에서 출력데이터의 크기를 입력데이터의 크기 그대로 유지하고, 풀링계층에서만 크기를 조절한다. 풀링에는 Max-Pooling과 Average



Pooling이 있는데 Max-Pooling은 해당영역에서 최대값을 찾는 방법이고, Average-Pooling은 해당영역의 평균값을 계산하는 방법이다. 이미지 인식 분야에서는 주로 Max-Pooling을 사용한다. 또한 풀링계층에서는 풀링의 윈도우 크기와 스트라이드 값은 같은 값으로 설정한다. 아래의 [그림 2-10]에서는 풀링의 윈도우 사이즈는 (2, 2)이며 스트라이드는 2로 설정하였다[6].



[그림 2-10] Max-Pooling 예제[6]

CNN이 최근 들어 다른 알고리즘에 비해 영상 분류 및 객체 검출에 우수한 성능을 보이는 이유는 크게 세 가지를 들 수 있다. 첫 번째는 Rectified Linear Unit(ReLU)이라는 활성화 함수(Activation Function)의 도입으로 이전 Sigmoid, tanh 등의 활성화 함수에서 나타나던 문제인 그레이디언트 베니싱(Gradient Vanishing) 문제가 없





어진 것이다. 그레이디언트 베니싱은 신경 회로망을 학습하는 대표적인 알고리즘인 오류 역전파 알고리즘에서 낮은 층으로 갈수록 전파되는 에러의 양이 적어짐으로 인해 그레이디언트 변화가 거의 없어져 학습이 일어나지 않는 현상이다. 이 문제로 인해 깊은 인공 신경망의 학습이 어려웠는데 ReLU의 도입으로 이 문제를 해결하여 깊은 인공 신경망에서도 낮은 층까지 학습이 가능해졌다. 두 번째 이유는 Dropout을 활용한 정규화를 들 수 있다. Dropout은 인공 신경망의 과적합을 방지하기 위해 학습 알고리즘 상에서 특정 비율의 뉴런을 무작위로 작동하지 않게 만든 채 학습을 수행하게 된다. 반복시마다 작동하지 않는 뉴런을 다르게 뽑아서 학습을 시켜 각각의 뉴런이 같은 정보를 학습하거나 아무런 정보도 학습하지 않는 것을 방지하였다. 위와 같은 이유로 CNN은 대용량의 영상 데이터가 존재할 때 영상 분류 및 객체 검출을 효과적으로 수행하며 현존하는 알고리즘 중 가장 좋은 성능을 보이는 것으로 보고되고 있다[5].





## 제3장 CNN을 이용한 모형 구축

### 3.1 연구의 대상

최근에는 환율을 예측하는데 단순한 선형 분석모형 보다는 비선형 분석 기법으로 환율을 예측하는 연구가 많이 진행되고 있다. 외환변화 패턴에 존재하는 비선형성은 국제금융분야에서 일반적 현상으로 받아들여지고 있기 때문에 다양한 비선형 형태를 탁월하게 인식할 수 있는 인공신경망을 환율예측모형으로 사용하는 연구가 활발히 진행되어왔고 더 나아가 딥러닝 기법을 이용하여 환율예측을 시도한다. 따라서 본 연구에서는 환율을 예측함에 있어서 CNN 모형을 사용하고자한다.

본 연구의 연구대상은 다음과 같다.

첫째, 2000년 이후의 환율자료를 이용하여 CNN 모형을 구축하여 원-달러 환율을 예측하여 보고 기존 신경망모형의 예측치와 비교해 어느 모형이 우수한 환율예측력을 보이는가를 비교해보는 것이다.

두 번째 목적은 신경망이론 중 CNN 알고리즘이 영상과 음성 외에 시계열 데이터에도 좋은 성능을 보일 수 있는지 연구해본다. CNN알고리즘의 장점이 시계열 데이터 예측에도 유효한지 알아보는 것이다.



### 3.2 자료의 수집

본 논문에서 사용하는 데이터는 한국은행 경제통계시스템을 통해 제공되는 공공데이터이며 2000년 1월3일부터 2016년 12월31일까지 기록된 데이터이다.

입력변수는 환율결정에 유의한 거시경제 변수들과 시계열 변수들을 선정하였다. 최종적으로 선정된 입력변수명과 변수설명은 [표 3-1]과 같다[7].

[표 3-1] 분석에 사용된 지표[7]

변수	변수설명
자본수지	경상 수지가 상품이나 서비스 거래를 대상으로 하는데 비해 자본수지는 정부와 민간이 해외로부터 차입등의 방식으로 외화를 도입하거나 이와는 반대로 해외에 신용공여 등의 방식으로 외화를 유출함으로써 발생하는 외화의 유·출입차를 나타낸다. 자본수지가 플러스라는 것은 외화의 유입이 유출보다 많았다는 것을 의미한다.
경상수지	①상품수지: 상품의 수출과 수입의 차액을 나타내는 수지. ②서비스수지: 해외여행, 유학·연수, 운수서비스 등과 같은 서비스 거래 관계가 있는 수입과 지출의 차액을 나타내는 수지. ③소득수지: 임금, 배당금, 이자처림 투자의 결과로 발생한 수입과 지급의 차액을 나타내는 수지. ④경상이전수지: 송금, 기부금, 정부의 무상원조 등 대가없이 주고 받은 거래의 차액을 나타냄



경제성장률	한나라에서 일정 기간(보통 1년) 동안 이룩한 국민경제(투자·산출량·국민소득)의 증감분을 전년도와 비교하여 산출해낸 비율.
물가	소비재 또는 생산재와 같이 특정의 범위에 속하는 상품의 가격을 종합·평균한 것이 소비재 물가 또는 생산재 물가
금리	미국연방준비제도이사회(FRB)에서 결정하는 연방기금금리. 미국의 금리가 인상이되면 달러가 미국내 투자수익의 증가로 달러가 우리나라 및 각국에서 유출되어 원-달러환율은 상승하게된다. 미국의 금리인상이 거듭되어 우리나라와 미국의 금리가 역전된다면 이런현상은 더욱 가중됨.
주가	주식의 시장가격. 이론적으로 배당청구권에 대한 가격, 즉 1주당 순자산지분을 표시해 배당금을 시장의 일반이자율로 자본환원해 결정하기도 하나, 실제로는 장래의 수익전망 경제동향 등을 감안한 주식시장의 수요공급에 의해 결정

각각 수집된 경제 지표와 환율데이터는 일별로 통합하여 하나의 원본데이터를 생성한다. 원/달러 환율자료는 여섯 개의 Input 변수를 선정하였으므로 6X6X1 CNN 행렬을 만들기 위해 월요일부터 토요일까지 총 6일치 일일 기준 환율자료를 수집하였다. 다만 데이터마다 일별로 기록되는 것이 있고 주 혹은 월별로 기록된 것이 있어 일



별로 데이터를 통일하기 위해 일별로 데이터를 만들었다. 데이터 전처리 방법은 뒤에서 좀더 자세히 설명하도록 하겠다.

[표 3-2]는 이렇게 구축된 원본 데이터(Source Data)의 예를 보인 것이다.

[표 3-2] 원본 데이터의 예

단위	금리	물가	자본수지	경제 성장률	경상수지	주가
변환	%	2015 =100	백만달러	%	백만달러	1980/01 =100
2000/01/03	N/A	65.83	-7.9	8.9	-217.6	943.88
2000/01/04	4.69	N/A	N/A	N/A	N/A	N/A
2000/01/05	4.71	N/A	N/A	N/A	N/A	N/A
2000/01/06	4.78	N/A	N/A	N/A	N/A	N/A
2000/01/07	4.77	N/A	N/A	N/A	N/A	N/A
2000/01/08	4.78	N/A	N/A	N/A	N/A	N/A

[표 3-3]은 모형을 구축하는 과정에서 전체 데이터를 분석용 (Traing) 데이터, 검증용(Test) 데이터 2 부분으로 분할하였는데 분석용 데이터는 모형구축에 사용되고 검증용 데이터는 모형의 예측을 측정하기 위한 목적으로 사용하였다.



[표 3-3] 데이터 분할 기간

	기간
Training Set	2000.1.3 ~ 2015.12.31
Test Set	2016.1.1 ~ 2016.12.31

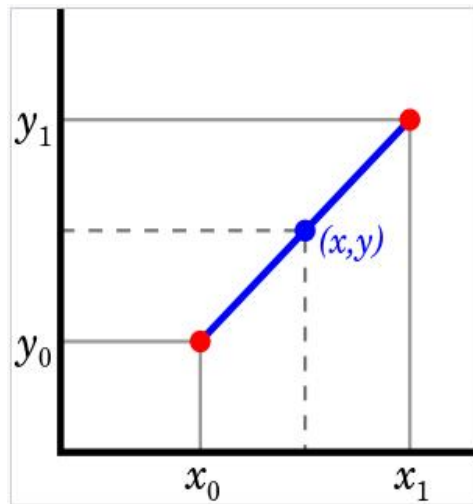
### 3.3 실험데이터 전처리

시계열 데이터들을 CNN 연산에 적용하기 위해서  $n \times n$ 으로 데이터를 처리해야했으므로 본 논문에서는 6개의 입력변수를 가로행으로 6일을 세로행으로 설정하여 6X6 Cell을 만들었다.

본 연구에서 사용하는 입력데이터는 수집기간단위가 달라 하나의 기준(일별)으로 맞추기 위해 선형보간법으로 데이터 전처리를 했다.

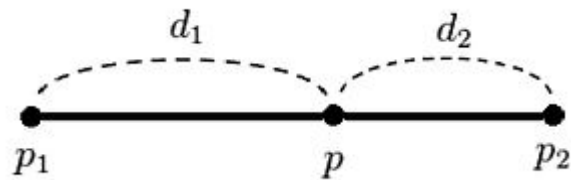
선형 보간법(Linear Interpolation)은 끝점의 값이 주어졌을 때 그 사이에 위치한 값을 추정하기 위하여 직선 거리에 따라 선형적으로 계산하는 방법으로 [그림 3-1]처럼 두 끝점  $(x_0, y_0)$  와  $(x_1, y_1)$  가 주어져 있을 때, 그 사이에 위치한  $(x, y)$  의 값을 추정하기 위해 두 점 사이에 직선을 긋고 다음과 같은 비례식을 구성할 수 있다.





[그림 3-1] 선형보간법 그래프

즉 [그림 3-2]처럼 일반화를 시켰을 때 선형보간법 그래프 두 지점  $p_1$ ,  $p_2$  에서의 데이터 값이 각각  $f(p_1), f(p_2)$  일 때,  $p_1, p_2$  사이의 임의의 지점  $p$ 에서의 데이터 값  $f(p)$ 는 [그림 3-3] 같이 계산할 수 있다.



[그림 3-2] 선형보간법 그래프(일반화)

단,  $d_1$ 은  $p$ 에서  $p_1$ 까지의 거리,  $d_2$ 은  $p$ 에서  $p_2$ 까지의 거리거리를 말한다.

이것을 수식으로 나타낸 것이 [그림 3-3]이다.

$$f(p) = \frac{d_2}{d_1 + d_2} f(p_1) + \frac{d_1}{d_1 + d_2} f(p_2)$$

[그림 3-3] 선형보간법 수식

만일 거리의 비를 합이 1이 되도록 정규화하면 ( $d_1 + d_2 = 1$ )  
[그림 3-3] 식은 [그림 3-4] 과 같이 단순화될 수 있다.

$$f(p) = d_2 f(p_1) + d_1 f(p_2)$$

[그림 3-4] 선형보간법 수식(정규화)

[표 3-4]는 [표 3-2]의 원본데이터에 선형보간법을 적용한 결과이다.



[표 3-4] 선형 보간법 적용 후 데이터

단위	금리	물가	자본수지	경제 성장률	경상수지	주가
변환	%	2015 =100(%)	백만달러	%	백만달러	1980 =100
2000/01/03	4.69	65.839	-7.9	8.9	-217.6	943.88
2000/01/04	4.69	65.847	-7.533	8.9	-223.342	939.07
2000/01/05	4.71	65.856	-7.167	8.9	-229.083	934.26
2000/01/06	4.78	65.864	-6.8	8.9	-234.825	929.44
2000/01/07	4.77	65.873	-6.433	8.9	-240.567	924.63
2000/01/08	4.78	65.881	-6.067	8.9	-246.308	919.82

또한, 신경망에서는 입력변수의 절대값이 크면 쉽게 포화상태에 빠지게 되므로 일반적으로 신경망의 데이터는 정규화 과정을 거치게 된다.

0~1 사이 범위의 값을 취할 때 최적의 성능을 제공하기 때문에 모든 입력변수와 목표변수를 0~1 사이로 조정하였다. 정규화단계에서는 데이터를 일정범위로 재조정하게 되는데, 그 범위는 신경망 입력의 최소범위에서 최대범위이다. 일반적으로 많이 사용되는 방법으로는 단순선형변화(Simple Linear Scaling)가 있고 그 식은 다음과 같다[9].





$$I = (D - D_{\text{mean}})/\text{STD};$$

I : 변환된 데이터,                      D : 실제 데이터

Dmean : 데이터의 평균값 , STD : 표준편차

위의 식을 적용하여 정규화 한 후의 Input데이터 값은 [표 3-5] 와 같다.

[표 3-5] 정규화 후 데이터의 예

단위	금리	물가	자본수지	경제 성장률	경상수지	주가
변환	%	2015 =100	백만달러	%	백만달러	1980 =100
2000/01/03	-1.74	-0.55	2.49	-0.91	-0.9	1.14
2000/01/04	-1.74	-0.51	2.49	-0.91	-0.91	1.14
2000/01/05	-1.76	-0.48	2.49	-0.91	-0.92	1.15
2000/01/06	-1.83	-0.44	2.49	-0.91	-0.93	1.21
2000/01/07	-1.83	-0.4	2.49	-0.91	-0.94	1.2
2000/01/08	-1.83	-0.36	2.49	-0.92	-0.94	1.21

선형보간법 사용 후 정규화한 데이터는 합성곱 계층에서 커널 적용 시 입력변수 편향을 줄이기 위해 입력변수들을 랜덤하게 일별로 배치하였다. [표 3-6]은 Input Data의 예를 표로 나타낸 것이다.



[표 3-6] Input Data

2000/01/03	-1.74	-0.55	2.49	-0.91	-0.9	1.14
2000/01/04	-0.51	2.49	-0.91	1.14	-0.91	-0.91
2000/01/05	-0.91	-0.92	1.15	-1.76	-0.48	2.49
2000/01/06	-0.91	1.21	2.49	-1.83	1.21	-0.44
2000/01/07	1.2	2.49	-0.4	-0.91	-0.94	-1.83
2000/01/08	-1.83	2.49	-0.36	-0.94	1.21	-0.92

또한 결과데이터(환율)은 9xx - 15xx 범위였는데, 이걸 0-65 범위의 Integer 로 매핑, 즉 9xx 을 베이스로 해서 10 단위로 예측하도록 하였다. 이는 나중에 Output 뉴런 개수를 66으로 선택한 것과 관련된다.

매핑 식은  $\text{Modified Out} = \text{Round}(\text{Out}/10) - 91$  으로 예로 1500원의 환율은 1500을 10으로 나눈 뒤 반올림 과정을 거쳐 91을 빼주기 때문에 값이 59로 매핑된다.



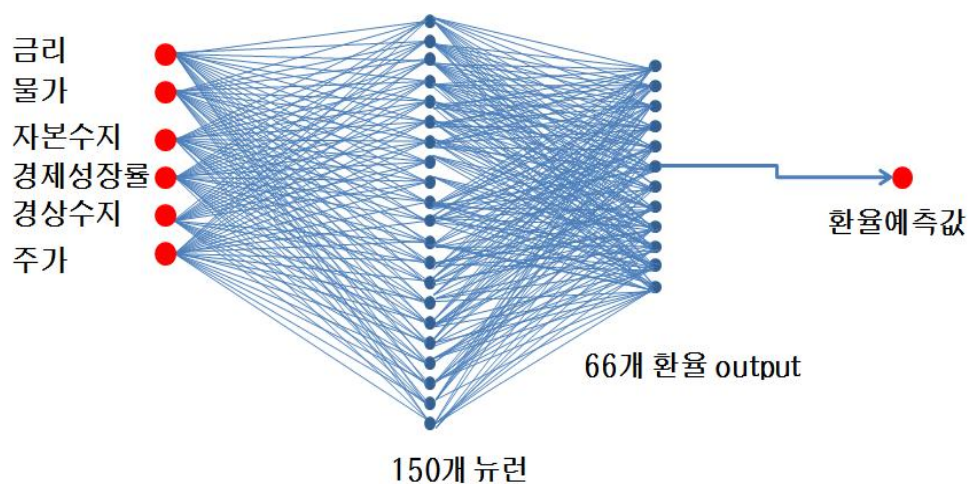
## 제4장 연구 결과 및 해석

### 4.1 선행 연구 모형 평가

#### 4.1.1 Multi Layer Perceptron 모형

CNN 모형에서 사용된 동일한 입력변수와 해석결과로 구축한 학습 자료를 이용해 다양한 알고리즘을 선택하여 인공신경망 학습이 가능한 MATLAB프로그램으로 학습을 실시하였다. 인공신경망의 최대 학습횟수는 10,000번으로 제한하였다.

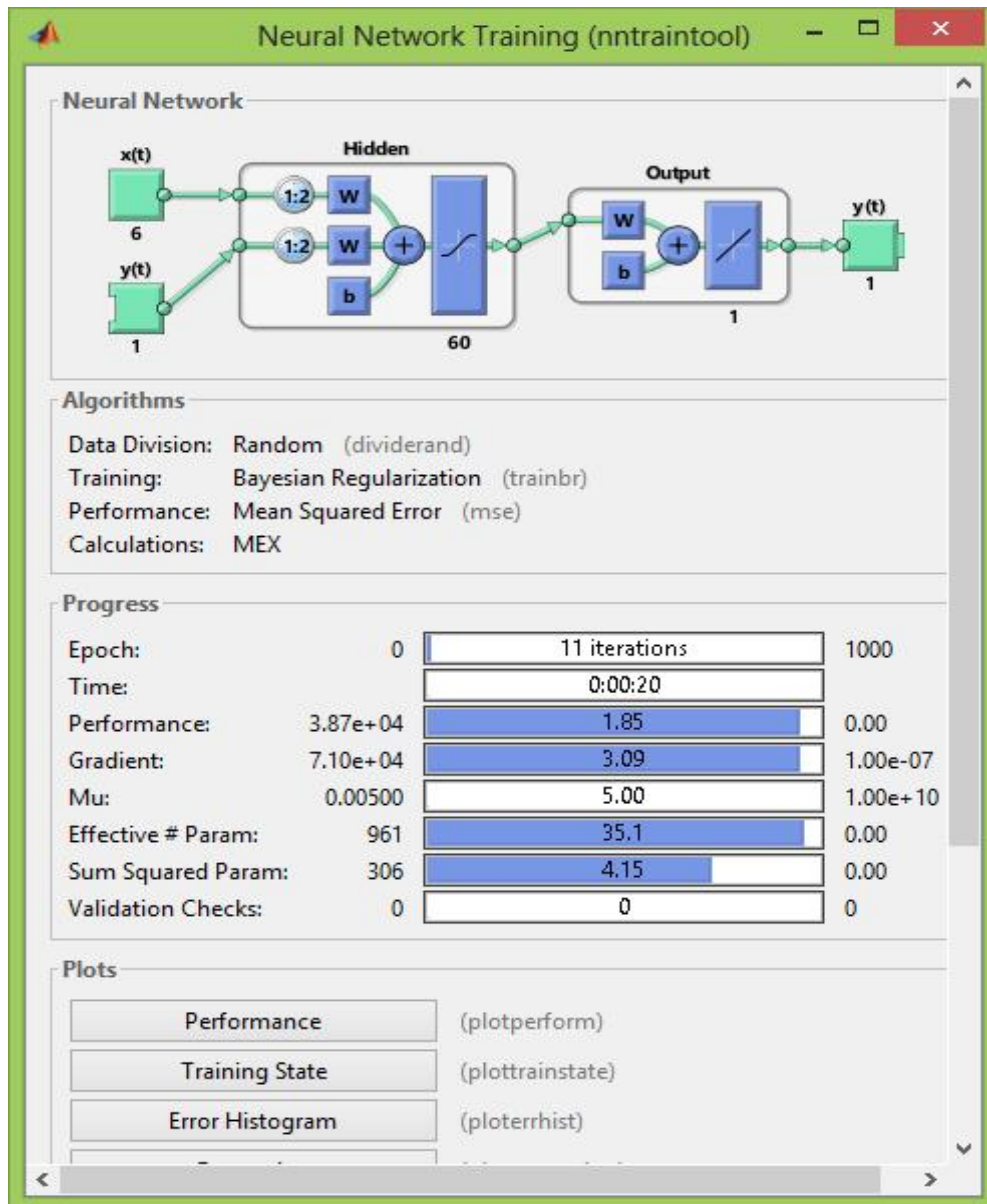
MLP 모형의 구조는 [그림 4-1]과 같이 1개의 은닉층으로 이루어져있다. 은닉층 노드 수 150개이고 66개의 범주내에서 환율예측값이 도출된다.



[그림 4-1] MLP 모형의 구조

#### 4.1.2 Multi Layer Perceptron 실험 결과

[그림 4-2]는 MATLAB을 사용하여 Training 및 Testing한 과정이다.



[그림 4-2] MLP 모델 MATLAB 실험 과정

[그림 4-2] 과정을 거쳐 [표 4-1] 결과를 도출하였다.

결과를 보면 평균제곱근오차(RMSE, Root Mean Square Error)가 1.72이며 오류검출율(MCR, Miss Classification Rate)은 0.687 (68%)이다.

[표 4-1] MLP 모델 실험 결과

	Samples	RMSE	MCR
Training Set	710	0.531225	0.435
Test Set	177	1.723659	0.687

## 4.2 구축 모형 평가

### 4.2.1 Convolutional Neural Network 모델

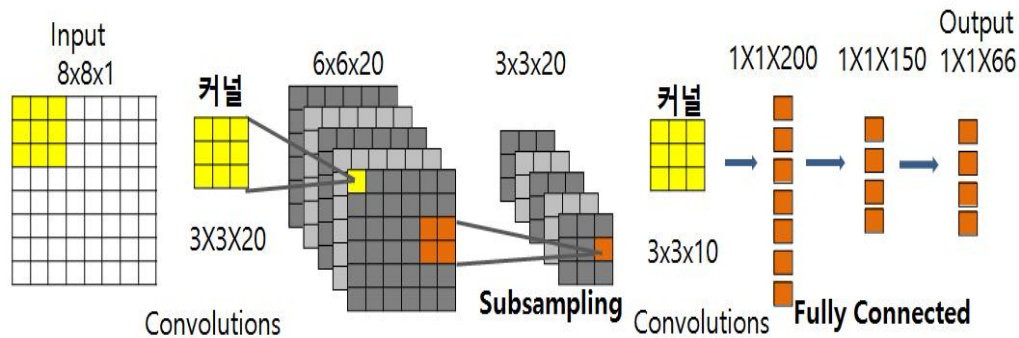
먼저 CNN 모델을 구현하고자 매개변수 연구를 통하여 적합한 커널수와 커널 사이즈, 뉴런 개수 등을 변경해가며 가장 좋은 성능을 보이는 매개변수를 선택하였다. [표 4-2]는 Convolution 2-Layer 적용하였을 때 결과이다.

[표 4-2] Convolution 2-Layer 적용시 매개변수 연구 결과

Model	1차 합성곱		2차 합성곱		뉴런수	결과	
	커널크기	커널수	커널크기	커널수		RMSE	MCR
A-1	3X3	10	3X3	10	150	0.25	0.61
A-2	3X3	10	3X3	15	150	0.25	0.65
A-3	3X3	15	3X3	15	150	0.27	0.56
A-4	3X3	15	3X3	20	150	0.25	0.61
A-5	3X3	20	3X3	15	150	0.29	0.57
A-6	3X3	20	3X3	10	150	0.24	0.55



연구 결과는 데이터 전처리를 통해 나온 Input값에 3X3 크기의 20개 커널을 적용한 컨볼루션 1계층과 3X3 크기의 10개 커널을 적용한 컨볼루션 2계층, Average 풀링 계층으로 이루어져 있는 2-layer CNN이며 ,컨볼루션 한 뒤 FC 계층은 은닉층이 없이 Softmax 출력 계층으로 뉴런 개수는 150 구조로 [그림 4-3]과 같다.



[그림 4-3] Convolution 2-Layer 적용시 CNN 최적 모델 구조

다음은 1-layer Convolution을 사용하여 실험을 하였다. [표 4-3]은 Convolution 1-Layer 적용하였을 때 결과이다.

1-layer Convolution 최적화 모델은 5X5 사이즈의 커널을 30개 적용여 합성곱을 한번 하고 뉴런수는 90으로 한 경우로 B-9 Model이다. RMSE 0.26 , MCR 0.57 으로 9개모델 중 가장 좋은 성능을 보였다.

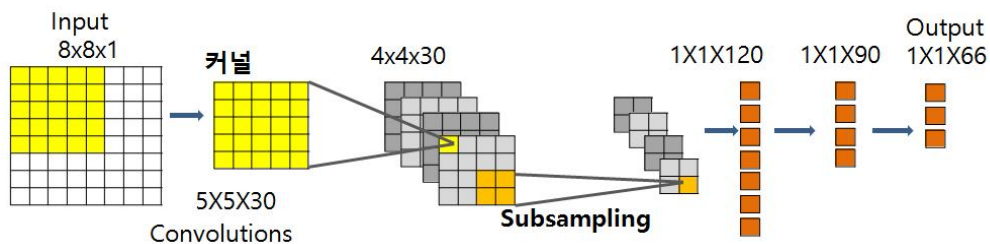


[표 4-3] Convolution 1-Layer 적용시 매개변수 연구 결과

Model	1차 합성곱		뉴런수	결과	
	커널 크기	커널 수		RMSE	MCR
B-1	3X3	10	120	0.31	0.67
B-2	3X3	20	120	0.4	0.76
B-3	3X3	30	120	0.29	0.7
B-4	3X3	20	60	0.31	0.7
B-5	3X3	20	180	0.34	0.71
B-6	3X3	20	240	0.38	0.71
B-7	5X5	10	50	0.26	0.79
B-8	5X5	20	70	0.33	0.82
B-9	5X5	30	90	0.26	0.57

#### 4.2.2 Convolutional Neural Network 실험 결과

앞서 Convolution 1-Layer 적용하였을 때 모델은 [그림 4-4] 와 같다.



[그림 4-4] Convolution 1-Layer 적용시 CNN 최적 모델 구조

연구 결과 1-Layer Convolution을 사용했을 때보다 2-Layer Convolution을 사용하고 커널 크기를 3보다 5로 하는 것이 더 성능이 좋은 것으로 판단된다.





### 4.3 실험 결과 분석

본 논문에서 실험 결과에 대한 성능 분석 및 평가는 학습결과의 평균 제곱근 오차(RMSE, Root Mean Squared Error), MCR(Miss Classification Rate)를 비교하여 학습과정의 적정성을 평가하였다. 평균 제곱근 오차(Root Mean Square Error; RMSE)는 추정 값 또는 모델이 예측한 값과 실제 환경에서 관찰되는 값의 차이를 다룰 때 흔히 사용하는 측도로 정밀도를 표현하는데 적합하다. 전체 평균과 각각의 정보가 얼마나 많이 차이가 나는가를 평가하는 것으로 RMSE는 작을수록 좋은 결과라고 생각할 수 있다

최종 성능은 [표 4-4]에서 MLP, CNN 1 Layer, 2Layer 세 개를 비교하였다. 기존 선행연구에서 사용되었던 MLP는 RMSE 1.73 이고 CNN은 RMSE 0.24 , MCR 0.55로 CNN이 MLP를 사용한 선행연구보다 좋은 성능을 보여주었다. 이 결과를 통해 시계열 데이터도 CNN을 통해 좋은 결과를 얻을 수 있음이 증명되었다.

[표 4-4] 성능 측정 방법별 성능

구분	RMSE	MCR
MLP	1.72	0.69
CNN (1Layer)	0.26	0.57
CNN (2Layer)	0.24	0.55





## 제5장 결론 및 향후 과제

### 5.1 결론

그동안 수많은 선행 연구에서 환율 예측 모델을 구축하였으나 기존의 통계적 또는 계량경제 모형이 주식가격, 환율 등과 같이 예측하기 어려운 비선형적인 특성을 가지는 시계열 자료의 분석에 일정한 한계를 보임에 따라 시계열 자료의 대안적 분석 모형으로 인공지능망이 제시되었고 많은 연구가 이루어졌다. 많은 은닉 층과 노드를 가진 인공지능망은 복잡한 문제의 학습을 통하여 예측이 가능하다.

하지만 너무 많은 뉴런의 개수는 과적합(Overfitting)의 위험이 있을 뿐만 아니라, 계산량도 상당히 많아서 고비용 발생의 원인이 된다[5]. 최근에는 이러한 인공지능망의 문제점을 보완하기 위해 딥러닝 기법이 다양하게 발달하고 있는 바 본 논문에서는 최근 이미지 인식, 음성 인식, 자연어 처리 등 다양한 분야에서 우수한 성능을 보여주고 있는 CNN 모형을 구축하여 환율을 예측하였다.

우선 환율 예측에 가장 적합한 CNN 모형을 찾기 위해 커널 개수, 크기, 뉴런 개수 변경하여 실험하였으며 영상과 음성 외에 시계열 데이터에도 좋은 성능을 보일 수 있는지 CNN의 성능을 평가하기 위해 선행 연구 모델인 MLP알고리즘과의 성능을 비교 하였다.

CNN 모델은 이미지, 영상에 적용이 용이하도록 만들어진 것으로 채널, 세로, 가로 이렇게 3차원으로 구성되므로 시계열 데이터를 3차



원으로 구성하는 방식을 구상하게 되었다.

자본수지, 경상수지, 경제성장률, 물가, 금리, 주가 6개의 입력변수를 사용하여 환율을 예측하고자 하므로 6X6X1 3차원으로 만들어주기 위해 가로는 6개의 경제지표(입력변수) 세로는 일자(6일)로 설정하였으며 합성곱 계층수(Convolution Layer)를 늘리기 위해 패딩(Padding)을 하였다.

CNN 모델의 매개변수 튜닝을 통해 성능을 향상시키기 위한 실험을 수행하였다.

결과적으로 기존 인경신경망 모델에 비해 CNN 모델로 환율을 예측하였을 때 13.9%의 성능 향상을 얻을 수 있었다. 이것은 환율과 같은 시계열 데이터도 CNN 모델로 예측에 효과가 있음이 증명되는 것이다.

최종적으로 본 연구를 통해 시계열 데이터 분석에 CNN 활용하는 것이 성능을 향상시키는데 효과적이라는 결론을 내리며, 향후 환율 예측에 빠른 훈련 및 예측 시간, LR/SVM에 비해 높은 분류 성능의 장점을 가진 CNN이 사용되는 것을 제안한다.



## 5.2 향후 과제

본 연구에서 사용한 입력변수는 선행 연구에서 인자값으로 사용하였던 기준과 동일하게 하였다. 시계열 데이터가 CNN모델에도 적합한지를 연구하는 것이 목적이었으므로 입력변수는 별도의 실험을 거치지 않고 선정하였다.

향후에는 환율예측에 영향이 있을 거시경제지표를 추가 발굴하여 입력변수로 선정하고 CNN모델 구조 변경 등을 통하여 성능 향상 관련 연구를 지속할 계획이다. 또한 RNN(Recurrent Neural Network) 및 LSTM(Long Short-Term Memory)과 같은 시계열 데이터에 사용되는 딥러닝 알고리즘과도 성능 비교하는 연구를 지속할 예정이다.



## 참고 문헌

- [1] 박상현, “신경망 기반 반도체 가격 예측 시스템”, 한양대학교 대학원 정보기술경영학과, 2010
- [2] Farzi, G. A., “An Image-Based Technique for Measuring Droplet Size Distribution: The Use of CNN Algorithm”, JOURNAL OF DISPERSION SCIENCE AND TECHNOLOGY, 2016
- [3] 김재현, “인공신경망모형과 ARIMA 모형의 원/달러 환율 예측성과 비교 연구”, 서강대학교 대학원 무역학과 석사학위논문, pp.5-81, 2002
- [4] 박종일, “시계열 데이터 전처리와 특징 선택을 활용한 기계 학습 기반의 효율적인 주가 방향성 예측 모델에 관한 연구,”서강대 정보통신대학원 석사논문, 2016
- [5] 김도우, “Doc2Vec을 활용한 CNN 기반 한국어 신문 기사 분류에 관한 연구”, 서강대 정보통신대학원 석사논문, 2016
- [6] 밑바닥부터 시작하는 딥러닝 파이썬으로 익히는 딥러닝 이론과 구현, 사이토 고키, 2017



- [7] 이경선, 신경망 기반 원-달러 환율예측 시스템, 한양대학교 정보통신대학원 석사논문, 2006
- [8] 신혜영, “미세먼지 농도 예측을 위한 통계 및 심층신경망 예측 기법 비교 연구”, 서강대 정보통신대학원 석사논문, 2016
- [9] 제대로 배우는 MATLAB, 방성완, 2014
- [10] 이지훈, “딥러닝을 이용한 주가 예측 모델”, 숭실대 대학원 미디어학과 석사학위논문, 2016
- [11] Wang, S.J., “A Path Planning Algorithm Based on the CNN Model under the Curvature Constraint”, Advanced Materials Research, 2014
- [12] Li, K., Thompson, S. “MLP Based Nonlinear Dynamic System Modeling Through Improved Training Algorithm”, Published for the International Federation of Automatic Control by Pergamon , 2002

