



저작자표시-변경금지 2.0 대한민국

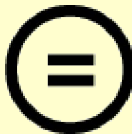
이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

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

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



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](#)

딥러닝 기법을 이용한
원/달러 환율예측 모형 제안
- DNN과 LSTM을 중심으로 -

연세대학교 일반대학원

투자정보공학 협동과정 전공

변 현 정

딥러닝 기법을 이용한
원/달러 환율예측 모형 제안
- DNN과 LSTM을 중심으로 -

연세대학교 일반대학원

투자정보공학 협동과정 전공

변 현 정

딥러닝 기법을 이용한
원/달러 환율예측 모형 제안
- DNN과 LSTM을 중심으로 -

지도교수 오 경 주

이 논문을 석사 학위논문으로 제출함

2020년 월 일

연세대학교 일반대학원

투자정보공학 협동과정 전공

변 현 정

변현정의 석사학위논문을 인준함

심사위원_____오 경 주 _____인

심사위원_____안 광 원 _____인

심사위원_____안 재 준 _____인

연세대학교 대학원

차 례

그림 차례	iii
표 차례	iv
국문 요약	v
제1장. 서론	1
1.1. 연구 배경	1
1.2. 연구 목적	2
1.3. 논문의 구성	3
제2장. 연구 배경	5
2.1. 인공지능(AI)를 이용한 환율예측 필요성	5
2.2. 환율예측에 적용한 DNN과 LSTM	6
2.3. 환율예측 성과 비교 연구	7
2.4. 슬라이딩 윈도우 기법 연구	9
2.5. 디노이징 오토인코더 연구	11
제3장. 연구 방법	12
3.1. 입력변수	12
3.2. DNN 모형	13
3.3. LSTM 모형	14

3.4. Denoising Autoencoder(DA) 모형	17
제4장. 실증 분석	19
4.1. 실증 분석	19
4.2. 슬라이딩 윈도우 분석	21
4.3. Denoising Autoencoder (DA) 슬라이딩 윈도우 분석	27
4.4. 추가분석(AR 시계열 분석)	33
제5장. 결론	39
5.1. 결론 및 시사점	39
5.2. 한계 및 향후 연구	40
참 고 문 헌	41
Abstract	47

그 립 차 례

Figure 1. Sliding window method	10
Figure 2. Exchange rate(close) Plot	12
Figure 3. DNN 구조	13
Figure 4. DNN Model	15
Figure 5. LSTM 구조	15
Figure 6. Autocorrelation 모형	17
Figure 7. DNN LSTM Model Predict Plot	19
Figure 8. Autoencoder DNN, LSTM Model Predict Plot	20
Figure 9. DNN, LSTM Model Predict Plot (2015)	22
Figure 10. DNN, LSTM Model Predict Plot (2016)	23
Figure 11. DNN, LSTM Model Predict Plot (2017)	24
Figure 12. DNN, LSTM Model Predict Plot (2018)	25
Figure 13. DNN, LSTM Model Predict Plot (2019)	26
Figure 14. Autoencoder DNN, LSTM Model Predict Plot (2015)	28
Figure 15. DA DNN, LSTM Model Predict Plot (2016)	29
Figure 16. DA DNN, LSTM Model Predict Plot (2017)	30
Figure 17. DA DNN, LSTM Model Predict Plot (2018)	31
Figure 18. DA DNN, LSTM Model Predict Plot (2019)	32
Figure 19. Autocorrelation Function Plot	34
Figure 20. AR1 Table	35
Figure 21. AR2 Table	36
Figure 22. AR3 Table	37

Figure 23. AR Model Predict Plot	38
--	----

표 차 례

Table 1. DNN Model	14
Table 2. LSTM Model	16
Table 3. RMSE, MAPE of DNN, LSTM Model	20
Table 4. RMSE, MAPE of Autoencoder DNN, LSTM Model	21
Table 5. RMSE, MAPE of DNN, LSTM Model (2015)	22
Table 6. RMSE, MAPE of DNN, LSTM Model (2016)	23
Table 7. RMSE, MAPE of DNN, LSTM Model (2017)	24
Table 8. RMSE, MAPE of DNN, LSTM Model (2018)	25
Table 9. RMSE, MAPE of DNN, LSTM Model (2019)	26
Table 10. Model Performance (RMSE)	27
Table 11. Model Performance (MAPE)	27
Table 12. RMSE, MAPE of DA DNN, LSTM Model (2015)	28
Table 13. RMSE, MAPE of DA DNN, LSTM Model (2016)	29
Table 14. RMSE, MAPE of DA DNN, LSTM Model (2017)	30
Table 15. RMSE, MAPE of DA DNN, LSTM Model (2018)	31
Table 16. RMSE, MAPE of DA DNN, LSTM Model (2019)	32
Table 17. DA Model Performance (RMSE)	33
Table 18. DA Model Performance (MAPE)	33
Table 19. AR Model RMSE, MAPE	38

국문 요약

딥러닝 기법을 이용한 원/달러 환율예측 모형 제안

- DNN과 LSTM을 중심으로 -

연세대학교 일반대학원

투자정보공학 협동과정

변 현 정

환율은 국가 간 무역에 지대한 영향을 미치는 매우 중요한 변수이다. 특히 외환리스크는 금융자산과 실물자산의 가치평가에 많은 영향을 미친다. 수출의존도가 높은 한국의 특성상 환 리스크의 예측은 매우 중요한 문제이다. 하지만 재무와 거시경제 분야에서 예측에 널리 쓰이던 기존 금융 시계열 모형의 한계가 드러났고, 이러한 상황에서 시계열 분석에 강점을 보이는 딥러닝 모형을 활용한 연구가 활발해지는 추세를 보인다. 하지만 국내에서 환율 예측에 대한 연구는 부족한 실정이다. 본 연구는 기존 국내 연구에서 한 번도 시도하지 않았던 DNN과 LSTM을 활용한 환율예측 모형을 제시한다.

본 연구에서 제시하는 환율예측 모형은 고시 환율의 시가, 고가, 저가, 종가, 거래량을 입력변수로 사용하였으며, 방대한 데이터를 다루려는 방법으로 DNN과 LSTM 모형을 사용하였다. 두 가지 모형의 실증분석에는 2010년부터 10년간의 데이터를 사용하였으며 슬라이딩 윈도우 기법을 이용하여 분석을 진행하

였다. 예측 모형의 성능을 더욱 향상시키기 위하여 본 연구의 목적은 디노이징 오토인코더(DA, Denoising Autoencoders)를 활용하였다. DA를 사용하여 데이터의 노이즈를 제거한 후 DNN과 LSTM의 시계열 예측 성능 향상을 기대한다.

본 논문에서는 전체 데이터에 대해 학습 기간을 7년으로 설정하여 모델을 최적화시켰다. 이후 설계한 두 가지 모형을 이용하여 환율예측을 수행하였고, DNN과 LSTM 두 모형의 예측 성능을 평가하는 지표로 RMSE와 MAPE를 활용하였다. 또한 딥러닝 모형 외에도 AR 모형을 벤치마크 모형으로 추가함으로 본 연구에서 제시하는 모형의 우수함을 입증하였다. 예측 결과 LSTM의 성능이 DNN의 성능보다 우수한 것으로 드러났다. 본 연구는 국내 환율예측 연구 중 LSTM을 활용한 최초의 연구이며, 향후 환율예측 연구에 실마리를 제공하는 데 의의가 있다.

주제어: 환율 예측, 슬라이딩 윈도우, 심층신경망, 장단기 기억 순환신경망, 디노이징 오토인코더, 자기회귀 모형

제1장. 서론

1.1. 연구 배경

1990년대 이후 우리나라 금융시장은 주식시장 개방, 금리자유화, 환율제도 변경 및 외환시장 개방화가 점진 진행되었다. 2019년 기준 GDP 대비 우리나라 수출입 의존도는 GDP 대비 수출입 비중이 크게 늘어나 70%에 육박할 정도로 주변국 대비 무역 의존도가 높은 것으로 나타났다. 이처럼 수출입 의존도가 높은 우리나라에서 원/달러 환율은 수출 물품의 가격 경쟁력에 매우 중요한 핵심 요소이며, 원자재와 같은 수입 물품 가격에 직접적인 영향을 미치고 있다. 또한, 환율은 한국 증권시장의 안정성을 유지하게 하고 있으며, 외국인 투자나 기관투자에 있어 투자 심리 안정을 유지하는 역할을 수행하는 변수로 작용하고 있다. 미국의 MSCI(Morgan Stanley Capital International)는 한국을 국제 금융시장에서 신흥 시장으로 분류하였는데, 주된 이유가 원화의 환전성이 부족하다는 것이다. 따라서 우리나라는 금융, 증권, 자본시장 환경이 다른 선진시장과 다소 상이하고 이에 따라 환율의 결정 또한 특수성이 존재하고 있다. 특히, 1997년 한국 IMF 외환위기 이후 환율예측 연구는 변동환율제로 인해 표본 확보의 어려움으로 매우 제한되어 있었다. 그에 따라 세계적으로 환율에 대한 예측은 활발하게 연구되고 있는 주제이지만 최근의 국내 원/달러 환율연구는 지극히 제한적으로 수행되었다.

국내 환율 연구는 고전적인 환율예측 모형을 이용하여 장단기 예측을 중심으로 이루어졌다(Park and Park, 2015)[31]. 특히 AI 기법을 이용한 주가 분석이 활발하게 이루어지고 있는 반면에 환율예측 연구는 고전 모형으로 수행되고 있는 것이 대부분이다. 환율에 대한 연구는 우리나라와 같은 수출입 중심, 무역 중심 국가에서는 필수적이고 매우 중요하지만, 딥러닝과 같은 최신 AI

기법으로는 연구되지 않고 있다. 따라서 본 논문은 AI 딥러닝 기법인 DNN(Deeper Neural Network)과 LSTM(Long Short-Term Memory)을 이용해 환율예측 연구를 수행하고 예측성능이 우수한 모형을 제시하고자 한다.

1.2. 연구 목적

AI 기술의 발달로 인하여 최근 학계에서는 AI를 활용한 재무변수 예측 연구가 활발히 진행하고 있다. 그중 본 연구에서 사용한 DNN과 LSTM 기법은 폭넓게 사용되고 있으며 유용성을 인정받는 분석 기법이다. 재무변수 예측에 쓰이는 DNN은 분류 기반 금융시장 예측연구(Dixon et al., 2017)[9]에 사용된 사례가 있다. 또한 재무 변수의 시계열적 특성이 중요한 경우 LSTM이 폭넓게 쓰이는데, 재무 시계열 특성 분석을 위한 Fischer et al.(2018)[11]의 LSTM 연구가 대표적이다. 또한, Cao(2020)[35]는 심층 LSTM을 이용하여 USD/CNY 환율예측 연구를 수행하였고, Cocianu and Avramescu(2020)[5]은 NARX모델에 LSTM을 결합하여 모형의 우수성을 입증하는 연구를 수행하였다. 이러한 연구 방법에 영향을 받아 우리나라에서도 딥러닝 기법을 이용한 재무 변수 예측 연구가 수행되었는데, 대표적으로 LSTM과 순환신경망 기법을 이용한 주가예측 연구(신동하 외 2명, 2017)[42]가 있다. 하지만, 국내 연구는 주가예측 연구를 중심으로 수행되고 있고, 아직 LSTM을 활용한 환율예측 연구가 수행되지는 않았다. 또한, LSTM이 기존 심층신경망에 비하여 주가예측 성능이 우수하다는 것을 보인 연구는 다수 존재하지만, 디노이징 오토인코더를 통해 시계열 데이터의 노이즈를 제거하고 예측 모형의 성능을 향상시키는 연구는 매우 미흡하다.

재무 변수들을 예측하고 분류하는 것에 AI 기술들의 유용성이 입증되고 있으며, 수출입 중심의 한국 금융환경에서는 환율예측 연구가 중요한 것이 분명하다. 그러나 한국의 환율예측 연구는 그 자체가 활발하게 이루어지지 않고

있으며 새로운 방법 또한 적용되지 않고 있는 것이 사실이다. 특히, AI 활성화 3대 요소로 데이터, 알고리즘, 컴퓨팅 파워가 급속도로 성장하면서 재무변수 분석 분야에도 AI 기법이 재조명받고 있으나 환율예측 연구는 적용한 연구가 없다. 따라서 본 연구는 AI 기술을 적용한 딥러닝 기법인 DNN과 LSTM을 중심으로 원(KRW)/달러(USD) 환율예측 연구를 수행하고자 한다.

본 연구는 환율 증가에 대한 예측 모형을 DNN과 LSTM을 이용하여 설계하였다. 모형의 입력 변수는 환율의 시가, 고가, 저가, 종가, 거래량이며, 슬라이딩 윈도우 기법(Sliding window method)을 활용하여 모형의 학습 기간과 예측 기간을 세분화하였다. 이 분석을 통하여 모형의 성능을 확인한 후 모형의 예측 성능을 향상하기 위하여 DNN과 LSTM 모형에 DA를 추가하여 성능을 비교한다. 최종적으로 고전적으로 사용되는 예측 기법인 AR 기법과 성능을 비교하여 본 논문에서 제안하는 모형의 우수성을 입증한다.

DNN 모형은 입력층에 5개의 노드, 중간층에 10개 노드로 3개 층, 출력층에 증가인 하나의 노드로 구성하였다. LSTM 모형은 입력층 5개의 노드를 가지며, 각 LSTM 셀의 히든 노드는 10개이고, 3개 층의 LSTM 층으로 구성하였다. 두 가지 모형의 활성화 함수는 ReLU로 설정하였고, 최적화 함수로 Adam Optimizer를 적용하여 연구를 수행하였다.

1.3. 논문의 구성

본 논문은 1장 서론에서 수출입 중심 국가인 한국의 특성에 따른 환율 연구의 필요성과 인공지능(AI) 기법을 이용한 환율예측 연구의 필요성에 대해 언급하였다.

2장에서는 환율예측에 대한 선행연구를 검토하고 DNN, LSTM을 환율예측과 주가예측에 활용한 선행연구 사례들을 제시하여 인공지능(AI) 기법의 환율예측에 적용이 가능한지 분석하였다. 3장에서는 실증 분석에 이용할 연

구 모형을 제안하며, 연구에 이용한 데이터의 특성, DNN과 LSTM, DA 모형의 특성을 구체적으로 서술하였다. 4장에서는 실증분석을 통하여 3장에서 제시한 모형을 상호 비교하여 환율예측 모형의 효용성을 제시하였다. 5장에서는 실증분석을 통해 얻은 결론과 본 연구의 시사점과 한계, 의의를 기술하였다.

제2장. 연구 배경

2.1. 인공지능(AI)를 이용한 환율예측 필요성

세계화 시대에 국가 간 교류가 활발해지면서 경제활동에 필요한 원자재는 물론 다양한 수출입 상품들이 거래되면서 환율은 국가 간 무역에도 큰 영향을 미치며 경제성장이나 경기회복에 지대한 영향을 미친다. 그에 따라 환율은 매우 중요한 재무변수로 작용하며 주식, 채권, 부동산 등 다양한 금융 및 실물 자산 가격에 영향을 미친다. 환율 변화는 재무변수의 상승과 하락에 영향을 미치는 주요한 변수이다. 특히 우리나라의 경우 수출의존도가 매우 높고, 국제시장에서 주도권을 행사하기에는 역부족인 소규모 개방경제에 해당하기 때문에 외환리스크에 매우 취약한 경제 구조를 이루고 있다. 이러한 상황에서 환 리스크 관리에 대한 연구는 필수적이며, 환율예측 연구는 더욱 중요하다.

환율예측은 전통적으로 계량경제 모형을 이용하여 이루어졌다. 계량경제 모형 중 대표적인 것으로 Engle의 ARCH(Autoregressive Conditional Heteroscedasticity: Engle, 1982)[10]모형이 있다. 하지만 계량경제모형은 시장에 존재하는 다양한 변수들의 변동성 문제로 인한 한계를 갖고 있다. Meese and Rogoff(1983)[29]는 거시 경제적 기초 요인을 고려한 랜덤워크 모델(Random Walk model)은 더 이상 일정 수준 이상으로 개선하는 것이 어렵다고 주장하였다. 이후 Meese와 Rogoff(1987)[29]는 통계적 환율예측 모형이 랜덤워크 모델보다 더 나은 성과를 얻을 수 없다는 연구 결과를 발표하며, 학계에 환율예측에 대한 ‘무용론’(無庸論)이 대두되기도 하였다. 이후에 랜덤워크 모형과 계량모형의 한계를 극복하기 위한 노력으로 Clarida와 Taylor(1997)[6]은 선물환 프리미엄의 기간 구조와 벡터 오차 수정모형을 활용해 개선된 예측력

을 제시하였으며, Kim et al.(2004)[24]은 동일한 방법으로 원/달러 예측 연구를 수행하였다. Kim(2000)[23]은 주 단위의 입력 변수와 데이터 및 월 단위 입력 변수로 환율을 예측하는 연구를 수행한 결과 기존의 계량경제모형보다 우월한 성과가 있었다. 하지만 계량 경제 모형의 한계인 정성적 요소를 변수로 활용할 수 없으며, 활용 가능한 자료의 크기가 상대적으로 작아서 Varian (2014)[33]은 가공되지 않은 방대한 양의 자료를 다룰 수 있는 강력한 도구의 필요성을 주장하였다.

2.2. 환율예측에 적용한 DNN과 LSTM 기법 특성

1989년 LeCun 등에 의해 소개된 DNN은 당시 신경망 학습에 소요되는 시간이 과도하다는 단점으로 인해 알고리즘이 뛰어남에도 불구하고 널리 사용되지 않았다(LeCun et al., 1989)[27]. 그러나 컴퓨팅 파워가 급격하게 향상됨에 따라 다시 DNN이 주목받고 있다. DNN(심층신경망, Deep Neural Network)은 입력층과 출력층 사이에 여러 층의 은닉층을 추가해 기존의 인공신경망보다 더 많은 층을 사용하고, 데이터에 특화된 표현과 분석 능력이 증가된 모형이다(Hinton et al, 2006)[19]. DNN은 비선형 관계를 모델링 하는데 용이하고 오류 역전파 알고리즘으로 학습되는 특징이 있다.

$$\Delta w_{ij}(t+1) = \Delta w_{ij}(t) - \varepsilon \partial C / (\partial w_{ij}(t+1))$$

LSTM(Long Short-Term Memory)은 순환신경망(RNN: Recurrent Neural Network)의 장기 의존성 문제를 극복하기 위해 제안된 특수한 형태의 순환신경망 모형이다. RNN은 입력값을 비선형 관계로 변환하는 활성화 함수인 하이퍼볼릭 탄젠트(Tanh)를 사용해서 출력을 낸다. 출력값은 -1과 1사이의 작은 값으로 결정되기 때문에 연산을 반복할수록 숫자가 점점 작아진다. 이

러한 절차가 반복되면 과거의 학습 결과가 소실되는 장기 의존성 문제(Gradient Vanishing Problem)가 발생하게 된다. 장기 의존성 문제가 발생하면 특정 단계부터 학습이 이루어지지 않는 문제가 발생하게 된다. 사라지는 과거의 가중치 문제를 극복하기 위해서 LSTM은 신경망 내부에 작은 메모리를 가진 셀(Cell)을 사용한다. 셀은 입력 게이트와 망각 게이트, 출력 게이트로 이루어져 있으며, 입력, 출력, 망각을 결정한다. 이 과정이 반복되면서 과거의 기울기가 소멸해 학습이 불가능한 일이 발생하지 않는다. LSTM은 기본적으로 순환신경망과 같은 방법으로 출력값을 계산하지만, 은닉층의 셀에서 입력, 망각, 출력 게이트를 적절하게 활용하여 정보 흐름을 조절한다. 이러한 결과로 LSTM 셀을 사용한 순환신경망은 금융 시계열 데이터와 같은 장기적인 시계열 학습이 필요한 데이터라도 기울기의 소실을 발생시키지 않는다.

2.3. 환율예측 성과 비교분석 연구

다음은 환율예측 결과 성과를 분석하는 방법에 대한 선행연구를 정리하였다. 김재현(2002)[40]은 시장 평균 환율 기간뿐만 아니라 변동환율 기간의 환율자료를 이용하여 인공신경망 모형을 구축하여 원/달러 환율을 예측하고 이를 시계열모형인 ARIMA 모형의 예측치와 비교해 어떤 모형이 우수한 환율예측성가를 보이는가를 실증적으로 비교·분석하였다. 본 논문에서는 세 구간으로 나누어 실험하였다. 첫 번째 실험구간은 1992년 7월1일부터 1994년 10월 31일까지이고 두 번째 실험구간은 1994년 11월1일부터 1996년 12월 31일까지이며 세 번째 실험구간은 1999년 1월1일부터 2000년 12월 31일까지의 세 가지 실험구간에서 적합한 환율예측 모형을 선정하여 마지막 90일의 환율을 예측하였으며, 이때 적용한 환율은 일일 원/달러 환율을 대상으로 하였다. 인공신경망 모형과 ARIMA 모형의 환율예측성가의 비교는 점예측 정확성 측면에서 평균 오차자승근과 평균 절대오차를 이용해 비교하였고, 변동 방향예측

정확성 측면에서 인공신경망 모형과 ARIMA 모형의 환율 변동 방향에 대한 적중률을 비교하였다. 그리고 두 환율예측모형의 예측력이 상이한 예측 기간 별로 서로 차이가 있는지를 통계적으로 검증하기 위해 두 모형의 예측 오차에 대한 회귀분석을 실시하였다. 분석 결과, 일일 원/달러 환율을 예측함에 있어 인공신경망 모형이 ARIMA 모형보다 우수한 예측성적을 보인다고 서술하였다[40].

Cocianu, C., & Avramescu, M. (2020)[5]은 재무 데이터를 예측하는 비선형 모델로 EURO와 US Dollar 사이의 환율을 예측하는 방법론에 대해 연구를 하였다. 예측 작업은 표준 NARX 신경망과 LSTM 기반 접근법의 두 가지 반복 신경망에 의해 수행되었다. 변수로 과거 거래 데이터와 주로 사용되는 기술 지표인 이동 평균 변형, 상위 볼린저 주파수 밴드 및 하위 볼린저 주파수 밴드(namely a variant of moving average, the Upper Bollinger Frequency Band and the Lower Bollinger Frequency Band.)로 구성되었다. 이들은 실험을 통해 LSTM 결합 NARX 모형의 우수성을 입증하였다[5].

김영철(2018)[39]의 연구에서는 외환위기 이후 한국이 자율변동환율제도로 전환되고 높아진 환율의 변동성으로 인해 환 리스크 관리가 매우 중요한 과제로 주목받자 이를 위해 환율예측의 필요성이 매우 높아졌다고 하였다. 본 연구는 인공신경망 모형인 다층퍼셉트론 방법론(MLP)을 이용해 원/달러 환율을 예측하였다. 환율예측을 위한 입력변수로는 현물환율, 원화 명목금리, 달러화 명목금리, 달러화 조달, 신용 가산금리로 설정하였으며 다층퍼셉트론 모형의 은닉층과 노드의 개수를 변화시키면서 실험을 진행해 환율예측에 가장 최적화된 구조를 제안하였고 다층퍼셉트론 모형의 예측력을 확인하였다[39].

박혜원(2019)[41]의 시계열모형의 예측에서는 LSTM 예측 성능에 관한 연구에서 고전적인 방법인 ARIMA와 LSTM의 예측 성능을 비교하였다. 연구 결과, 모형 식별이 정확한 경우에는 모수적 방법이 LSTM(ADAM)과 성능이

비슷하고, LSTM(SGD)보다는 성능이 우수하였다. 또한, 모형 식별 오류가 발생한 경우에는 LSTM(ADAM)의 성능이 모수적 방법보다 많은 경우가 우수하다고 하였다[41].

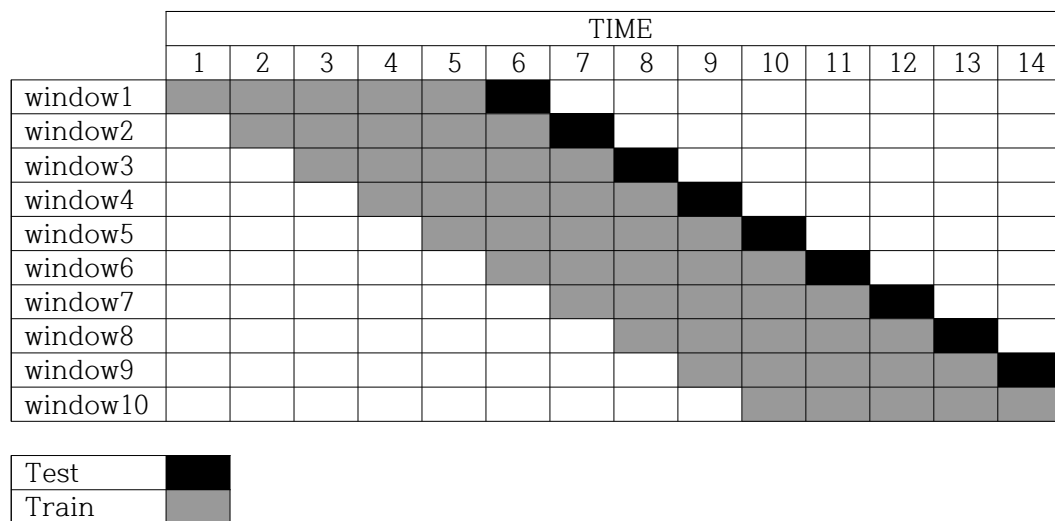
강민영(2017)[38]의 연구에서는 환율예측 모형을 구축했으나 통계적 또는 계량경제 모형이 주식가격, 환율 등과 같이 예측하기 어려운 비선형적인 특성을 가지는 시계열 자료의 분석에 일정한 한계를 보임에 따라 시계열 자료의 대안적 분석모형으로 인공신경망을 제시하였다. 최근에는 인공신경망의 문제점을 보완하기 위해 딥러닝 기법이 다양하게 발달하고 있다. 강민영(2017)은 비정형적인 분야에서 우수한 성능을 보여주는 합성곱 신경망의 환율예측을 적용해 시계열 데이터 성능을 평가하는 것을 연구 목적으로 하고, 그 결과 환율의 흐름 파악이 가능하고, 인공신경망과 관련된 다른 많은 연구와 비교했을 때 매우 좋은 성능을 보였다[38].

이낙영(2019)[43]의 연구는 KOSPI200 선물지수 예측에 디노이징 필터를 사용하여 LSTM의 모델 예측 성능을 향상 시키는 것이다. 슬라이딩 윈도우 실험을 통해 디노이징 필터를 적용한 LSTM의 성능이 향상되었고 디노이징 필터 중 사비츠키-골레이 필터(Savitzky-Golay Filter)가 모델 예측성능 향상에 유용함을 입증하였다. 본 연구는 2007년부터 12년간 데이터와 2017년부터 2년간의 KOSPI200 선물지수 시, 고, 저, 증가 및 거래량을 사용하여 노이즈 제거 후 LSTM의 입력변수로 사용하였다. 슬라이딩 윈도우 실험에서는 일봉 training 5년 test 1년 30분봉 training 1년 test 1개월로 진행하였고 예측 모형 평가는 RMSE와 MAPE로 평가하였다[43].

2.4. 슬라이딩 윈도우 기법 연구

본 논문은 제시한 모형의 적합성을 검증하고 오차율을 확인하기 위해 학습 구간과 예측구간을 유동성 있게 설정하는 모형인 슬라이딩 윈도우 기법(Sliding Window Method)

ng window method)을 사용하였다. 슬라이딩 윈도우 기법은 입력변수를 만
 들고, 입력 변수와 출력값(예측값)을 구간별로 누적 시켜 학습시키는 특징을
 가지고 있다. 이번 연구에서는 2010년부터 2019년까지의 환율 데이터에 슬라이
 딩 윈도우 기법을 이용하여 초기 5일의 환율 데이터로 익일의 환율 증가를
 예측하도록 설계했다. 슬라이딩 윈도우 기법은 시계열성 특징을 가지는 데이
 터를 분석하는데 매우 적합한 기법이다. 학습 기간으로 설정한 기간의 데이
 터는 일회성으로 활용되지 않고, 1~5년의 데이터는 6년의 예측, 2~6년의 데
 이터는 7년의 예측에 사용되어 입력데이터의 손실을 최소화하고, 입력 변수
 의 시계열성을 유지해줄 수 있는 장점을 가지고 있다. 그 때문에 슬라이딩
 윈도우 기법은 금융 시계열 분석과 예측에 유용한 방법으로 알려져 있다(Ref
 enes et al., 1993; Cheong과 Oh, 2014)[30,7].



<Figure 1. Sliding window method>

본 논문에서는 <Figure 1>과 같이 슬라이딩 윈도우(Sliding window)를 적용해
 서 5년의 학습과 1년의 실험으로 학습 기간을 충분히 확보해 높은 수준의 환

율예측 연구를 실시한다. 슬라이딩 윈도우 기법(Sliding window method)의 기대 효과로 상대적으로 작은 환율 데이터 크기의 한계를 극복할 수 있을 것으로 판단된다.

2.5. 디노이징 오토인코더(DA, Denoising Autoencoder)

앞서 환율예측 성과 비교분석 연구에서 디노이징 오토인코더를 이용하여 데이터를 필터링한 LSTM 모형의 시계열 데이터 예측 성능이 기존의 LSTM 보다 향상되었다는 것이 이전 연구(이낙영, 오경주, 2019)[43]를 통해 확인되었다. Lin(2004)[28]은 오토인코더 학습에 노이즈가 가미된 입력으로부터 원래의 입력을 출력하도록 학습하는 디노이징 오토 인코더(Denoising autoencoder)를 제안하였으며 실험을 통해 입력에 작은 변화(노이즈)가 있어도 최적화된 실측값을 얻을 수 있음을 보였다. Vincent, p., et al.(2008)[34]에서는 최적화된 실측값을 얻기 위한 방법의 하나로 시퀀셜 디노이징 오토인코더(SDAE)를 제안하였으며 노이징 함수로 삭제와 순서 변경을 사용하였다. 해당 논문 또한 실험을 통해 디노이징 메커니즘을 이용하여 최적화된 실측값을 생성할 수 있음을 보였다.

제3장. 연구 방법

3.1. 입력변수

본 연구에서는 2010년 1월 4일부터 2019년 12월 11일까지의 일별 원/달러 환율의 시가, 고가, 저가, 종가, 거래량 데이터를 활용해서 연구를 수행하였다. 장기기억 모델을 학습시키기 위해서 10년 치 환율 데이터를 확보하였다. 총 2,454일의 기간 동안의 원/달러 환율 데이터를 사용하였고, 학습과 예측에 사용된 데이터의 비율은 7:3으로 설정하여 다음 <Figure 2>와 같이 연구를 수행하였다.



<Figure 2. Exchange rate(close) Plot>

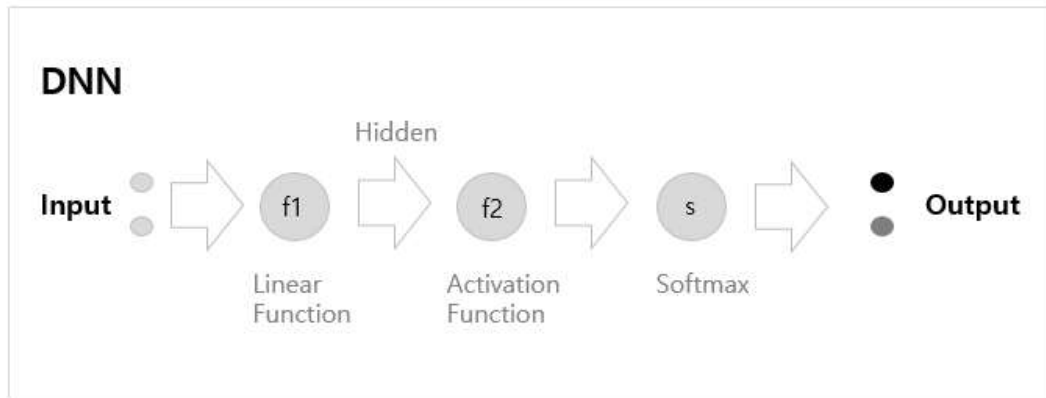
DNN과 LSTM 모형의 예측 성능에 대한 평가로서 평균 제곱근편차(RMSE: Root Mean Squared Error)와 평균 절대백분율오차(MAPE: Mean Absolute Percentage Error)를 활용하였다. RMSE는 실제 데이터값과 예측한 데이터의 값의 차이를 다루는 데 사용하는 척도이고, 특히 모형의 정밀도, 정확도를 표현하는 데 적합하다. MAPE는 오차가 예측값에서 차지하는 정도를 나타내는

지표로서 모델 신뢰성 검증에 가장 일반적으로 사용된다. 실제값이 크면 오차도 커지며 실제값이 0에 가까워지면 오차의 영향력이 무한대로 커져 계산이 불가능하다. 따라서 RMSE 값과 MAPE 값이 작을수록 예측모형이 정밀하다고 할 수 있다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (real_t - predict_t)^2} \quad , \quad MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{predict_t - real_t}{predict_t} \right|$$

3.2. DNN 모형

심층신경망인 DNN은 입력층과 출력층 사이에 2개 이상의 은닉층으로 이루어진 인공신경망으로서 비선형 관계를 모델링 하는데 용이하며 오류 역전파 알고리즘으로 학습이 된다(G.E Hinton et al, 2006)[19].



<Figure 3. DNN 구조>

$$\Delta w_{ij}(t+1) = \Delta w_{ij}(t) - \varepsilon \frac{\partial c}{\partial w_{ij}(t+1)}$$

Δw_{ij} : visible unit i 와 hidden unit j 의 가중치의 변화량

ϵ : 학습률

c : 비용함수

t : 데이터 시점

본 연구에서 DNN 모형은 다음 <Table 1>와 같이 설정하였다. 입력층 5노드, 중간층 10노드로 하여 3층, 출력층 1노드로 구성했다. 입력 데이터의 개수(시가, 고가, 저가, 종가, 거래량)에 맞추어 입력 노드를 5개로 사용하였고, 중간층(은닉층)의 개수는 시행 착오법을 이용하여 결정하였다. 오차감소속도가 빠른 ADAM을 최적화 도구로 사용하였고, 당일의 변수를 활용해서 내일의 종가를 예측하는 모형으로 설계하였다.

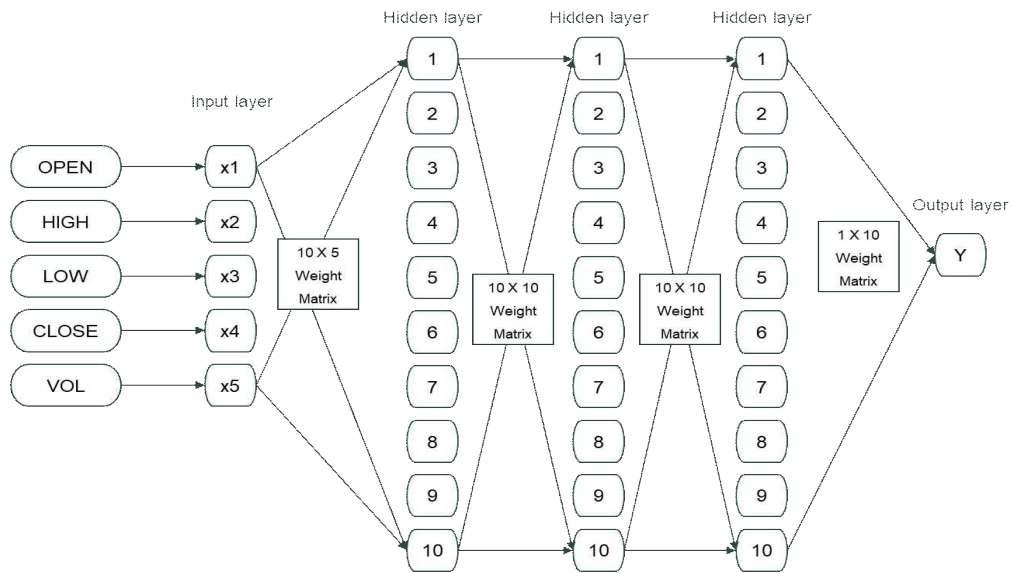
<Table 1. DNN Model>

항 목	내 용
Layer	10개의 뉴런으로 된 3개 Hidden Layer
활성화 함수	ReLU(Rectified Linear Unit)
최적화 도구	Adam Optimizer
입력층	시가 / 고가 / 저가 / 종가 / 거래량
출력층	내일의 종가

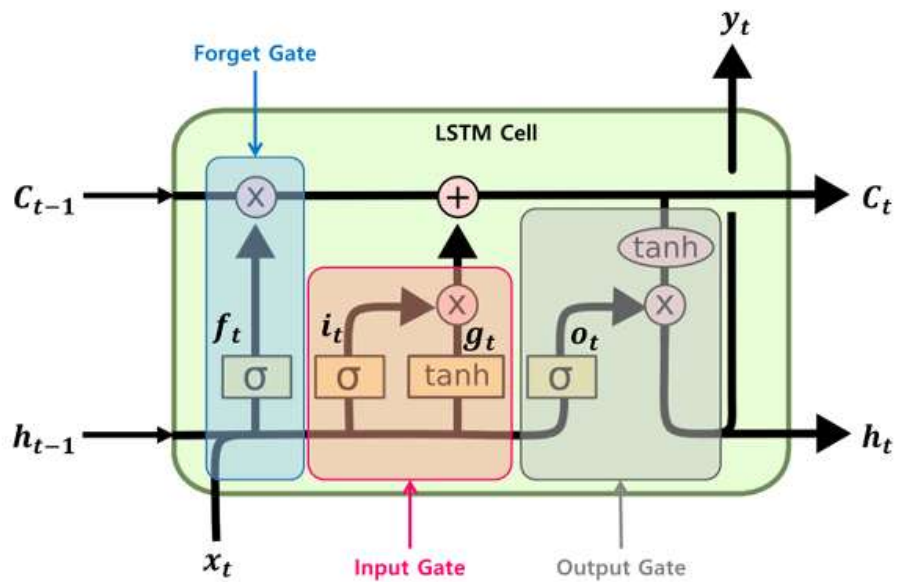
3.3. LSTM 모형

LSTM은 RNN(순환신경망) 구조의 많은 변형된 모델 중 하나이다. RNN의 Vanishing gradient(기울기 소실) problem과 느린 학습 속도를 해결하기 위해 고안된 모델(Hochreiter and Schmidhuber, 1997)이며 합 연산 구조를 가지고 있어 기울기가 소실되지 않으며 RNN에 비해 긴 시퀀스 입력을 처리 할 수

있다. LSTM 구조는 입력게이트, 망각게이트, 출력게이트, 단기상태, 장기상태로 이루어져 있다(Gers et al, 1999)[15].



<Figure 4. DNN Model>



<Figure 5. LSTM 구조> [15]

$$\begin{aligned}
 & \cdot f_t = \sigma(W_{xh_f}x_t + W_{hh_f}h_{t-1} + b_{h_f}) \text{ -망각게이트} \\
 & \cdot i_t = \sigma(W_{xh_i}x_t + W_{hh_i}h_{t-1} + b_{h_i}) \text{ -입력게이트} \\
 & \cdot g_t = \tanh(W_{xh_g}x_t + W_{hh_g}h_{t-1} + b_{h_g}) \\
 & \cdot o_t = \sigma(W_{xh_o}x_t + W_{hh_o}h_{t-1} + b_{h_o}) \text{ -출력게이트} \\
 & \cdot c_t = f_t \times c_{t-1} + i_t \times g_t \text{ - 망각게이트} \\
 & \cdot h_t = o_t \times \tanh(c_t) \text{ - 출력정보}
 \end{aligned}$$

σ - 시그모이드함수

W_{x_o} 는 x_t 와 함께 각 게이트에서 사용되는 가중치

W_{h_o} 는 h_{t-1} 와 함께 각 게이트에서 사용되는 가중치

b_o 는 각 게이트에서 사용되는 편향

이 연구에서는 LSTM 모델은 입력 층 5노드, 시퀀스 길이 7, 각 셀의 히든 노드는 10개로 설정하였다. 그리고 3층의 LSTM을 쌓은 형태의 모형을 구축했다. 입력층의 노드는 시가, 고가, 저가, 종가, 거래량의 데이터를 사용하여 5개의 노드로 구성하였고, 시퀀스의 길이와 은닉층의 결정은 시행 착오법을 통하여 결정하였다. 오차감소속도에서 우위를 보이는 ADAM을 최적화 도구로 사용했고, 당일의 변수를 이용하여 익일의 종가를 예측하는 모형으로 설계하였다. 다음의 표 <Table 2> 는 모형의 개요를 나타낸다.

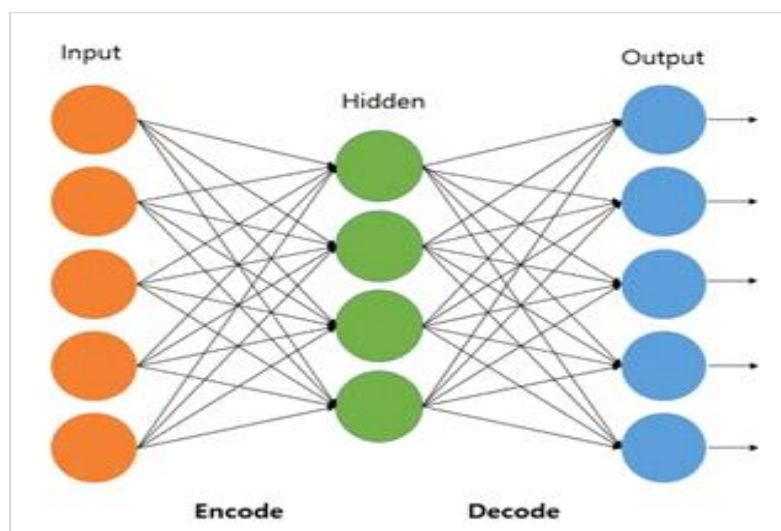
<Table 2. LSTM Model>

항 목	내 용
Layer	입력층 5노드
	시퀀스 길이 7
	각 셀의 히든노드 10개
	Stacked LSTM Layers 3개
활성화 함수	ReLU(Rectified Linear Unit)
최적화 도구	Adam Optimizer

입력층	시가 / 고가 / 저가 / 종가 / 거래량
출력층	내일의 종가

3.4. Denoising Autoencoder (DA)모형

대표적 비지도 학습인 Autoencoder(오토인코더)는 아주 간단한 뉴럴 네트워크 레이어의 재구성이다. 아래 <Figure 6>와 같이, 히든 레이어(Hidden Layer) 하나를 가지는 모양의 뉴럴 네트워크이다. Input data와 Output data가 동일한 신경망 학습 방법으로 비지도 학습이나 자기 지도학습에 이용된다. Autoencoder는 입력값과 히든레이어의 가중치를 계산해 함수를 통과시킨 후 결과물과 출력 레이어 값의 가중치를 계산해 다시 함수를 통과시킨다. 이렇게 나온 출력값을 이용해 MSE(Mean Squared Error)를 계산하고 이 값에 loss 값을 SGD(Stochastic Gradient Descent)로 최적화시킨 후 오류 역전파를 사용하여 가중치를 갱신한다.



<Figure 6. Autoencoder 모형> [14]

디노이징 오토인코더는 입력층과 히든 레이어 사이에 노이즈를 추가하고, 노이즈가 없는 원본 입력을 재구성하도록 학습시킴으로써 의미 있는 특성을 학습하도록 제약을 주는 학습 방법이다[14]. 이 연구에서 오버피팅을 줄이기 위해 인코더와 디코더를 사용하였고, 시퀀스 길이 7 최적화 도구로 Adam Optimizer를 사용하였다.

제4장. 실증 분석

4.1. 실증 분석

이 연구는 전체 학습 기간(2010/01/04 ~ 2019/12/11)인 2,454일 중 70%인 1,721일을 학습데이터로, 30%인 738개의 데이터를 실험 데이터로 설정해 분석하였다. 다음 <Figure 7>은 전체 기간의 학습에 대한 DNN과 LSTM 모형의 예측 추세를 나타낸 그림이다. 학습 기간이 장기화되면 부정확한 결과가 나올 가능성이 있기 때문에 2010년의 환율 데이터를 기점으로 학습을 진행하였다. 분석에 사용한 프로그램은 Python이며 LSTM 실험을 위한 모듈은 Pytorch, Keras을 사용하였다.



<Figure 7. DNN, LSTM Model Predict Plot>

<Table 3. RMSE, MAPE of DNN, LSTM Model>

	LSTM	DNN
RMSE	0.01975872	0.02234
MAPE	3.73122127	3.654525

그래프의 X축 데이터는 각각의 실험데이터 일자를 일련번호로 나타낸 값이고, 실제 값은 2016년 12월 8일부터 2019년 12월 11일까지의 일자를 의미한다. 파란색으로 나타낸 선이 환율의 증가이고, 초록색 선은 LSTM 모형 나타내며 붉은 색 선은 DNN 모형을 통해 예측한 환율 값을 나타낸다. LSTM과 DNN 모형의 예측치와 실제 환율의 증가를 비교하여 RMSE 와 MAPE를 산출하였다. 에포크와 배치사이즈는 각각 20000으로 세팅하였고 각각의 값은 <Figure 8>과 같다.



<Figure 8. Autoencoder DNN, LSTM Model Predict Plot>

<Table 4. RMSE, MAPE of Autoencoder DNN, LSTM Model>

	LSTM	DNN
RMSE	0.005016	0.050027
MAPE	3.516912	3.363806

4.2. 슬라이딩 윈도우 분석

이 연구에서는 학습데이터와 실험데이터는 일정 간격으로 학습 데이터 5년, 실험 데이터 1년으로 분할하고, 실험데이터의 크기만큼 이동시키며 실험을 진행하였다. 시계열 자료의 경우 과거의 자료가 분석에 지나친 악영향을 줄 우려가 있어 데이터 일부만 학습데이터로 사용하였다. 슬라이딩 윈도우로 시계열 자료를 분석할 경우 과거 데이터 영향을 최소화 할 수 있음이 밝혀졌다(Kim and Oh, 2019)[26]. 슬라이딩 윈도우를 통한 분석은 총 5번에 나누어 진행하였으며, 분석의 정확도는 RMSE와 MAPE를 이용하여 평가하였다.



<Figure 9. DNN, LSTM Model Predict Plot (2015)>

<Table 5. RMSE, MAPE of DNN, LSTM Model (2015)>

	LSTM	DNN
RMSE	0.0259633	0.04849574
MAPE	3.6726605	3.55928324

2015년 예측에 사용된 학습데이터는 <Figure 9>와 같이 2010년 1월 4일부터 2014년 12월 30일까지의 데이터이며, 실험 데이터는 2015년 1월 2일부터 2015년 12월 30일까지의 데이터이다.



<Figure 10. DNN, LSTM Model Predict Plot (2016)>

<Table 6. RMSE, MAPE of DNN, LSTM Model (2016)>

	LSTM	DNN
RMSE	0.05282965	0.0715836909
MAPE	3.26403976	3.4181700317

2016년 예측에 사용된 학습데이터는 <Figure 10>와 같이 2011년 1월 3일부터 2015년 12월 30일까지의 데이터이며, 실험 데이터는 2016년 1월 3일부터 2016년 12월 29일까지의 데이터이다. 해당 기간에서 RMSE 수치가 다소 상승하였는데 이는, 해당 기간의 환율 변동성이 다른 기간에 비하여 높고 실제로 환율의 변화가 불규칙하게 일어났다고 해석할 수 있다.



<Figure 11. DNN, LSTM Model Predict Plot (2017)>

<Table 7. RMSE, MAPE of DNN, LSTM Model (2017)>

	LSTM	DNN
RMSE	0.0259633	0.04849574
MAPE	3.6726605	3.55928324

2017년 예측에 사용된 학습데이터는 <Figure 11>와 같이 2012년 1월 4일부터 2016년 12월 29일까지의 데이터이며, 실험 데이터는 2017년 1월 2일부터 2017년 12월 28일까지의 데이터이다.



<Figure 12. DNN, LSTM Model Predict Plot (2018)>

<Table 8. RMSE, MAPE of DNN, LSTM Model (2018)>

	LSTM	DNN
RMSE	0.022251115	0.0372352410
MAPE	2.737501534	2.6318155329

2018년 예측에 사용된 학습데이터는 <Figure 12>와 같이 2013년 1월 2일부터 2017년 12월 28일까지의 데이터이며, 실험 데이터는 2018년 1월 2일부터 2018년 12월 28일까지의 데이터이다.



<Figure 13. DNN, LSTM Model Predict Plot (2019)>

<Table 9. RMSE, MAPE of DNN, LSTM Model (2019)>

	LSTM	DNN
RMSE	0.0230981	0.0738295
MAPE	2.8855749	3.0163454

2019년 예측에 사용된 학습데이터는 <Figure 13>와 같이 2014년 1월 2일부터 2018년 12월 28일까지의 데이터이며, 실험 데이터는 2019년 1월 2일부터 2019년 12월 11일까지의 데이터이다.

다음 표 <Table 10, 11>에 DNN과 LSTM 모형의 RMSE 값과 MAPE 값을 예측 기간별로 구분하여 표시하였다.

<Table 10. Model Performance (RMSE)>

	전체 기간	15년 예측	16년 예측	17년 예측	18년 예측	19년 예측
DNN	0.02234	0.048495	0.0715836	0.0508445	0.037235	0.07382
LSTM	0.02060	0.025963	0.052829	0.0242966	0.022251	0.02309

<Table 11. Model Performance (MAPE)>

	전체 기간	15년 예측	16년 예측	17년 예측	18년 예측	19년 예측
DNN	3.654525	3.559283	3.418170	2.3661267	2.6318155	3.0163454
LSTM	3.709835	3.672660	3.264039	2.3004927	2.7375015	2.885574

LSTM모형의 RMSE과 MAPE는 다음과 같이 관찰되었다. 2015년 예측값과 실제 값의 RMSE는 0.025963, MAPE는 3.672660으로 나타났고, 2016년의 RMSE는 0.052829로 MAPE는 3.264039로 나타났다. 이어 2017년에는 RMSE는 0.0242966, MAPE는 2.7375015로 나타났고, 2018년에는 RMSE는 0.022251, MAPE는 2.7375015로 나타났다. 2019년에는 RMSE는 0.02309, MAPE는 2.885574의 값을 갖는 것으로 확인되었다. 따라서 RMSE예측값을 통해 LSTM 예측력이 DNN보다 전 구간에서 우수하고 MAPE에서는 일부구간에서 우수함을 보였다.

4.3. Denoising Autoencoder (DA) 슬라이딩도어 분석

실험데이터는 2003년부터 2020년으로 학습데이터와 실험데이터를 일정 간격으로 학습 데이터 7년, 실험 데이터 3년으로 분할하여 2010년부터 2020까지 디노이징 데이터를 추출한 후 학습데이터와 실험데이터를 일정 간격으로 학습

데이터 5년, 실험 데이터 1년으로 분할하고 실험데이터의 크기만큼 이동시키며 실험을 진행하였다. 시계열 자료의 경우 과거의 자료가 분석에 지나친 악영향을 줄 우려가 있어 데이터의 일부만 학습데이터로 사용해 진행하였다. 슬라이딩 윈도우를 통한 분석은 총 5번 진행하였으며, 정확도는 RMSE와 MAPE를 이용해 평가하였다.



<Figure 14. Autoencoder DNN, LSTM Model Predict Plot (2015)>

<Table 12. RMSE, MAPE of DA DNN, LSTM Model (2015)>

	LSTM	DNN
RMSE	0.00595	0.04432
MAPE	3.83962	3.59286

2015년 예측에 사용된 학습데이터는 <Figure 14>와 같이 2010년 1월 4일부터 2014년 12월 30일까지의 데이터이며, 실험 데이터는 2015년 1월 2일부터 2015년 12월 30일까지의 데이터이다.



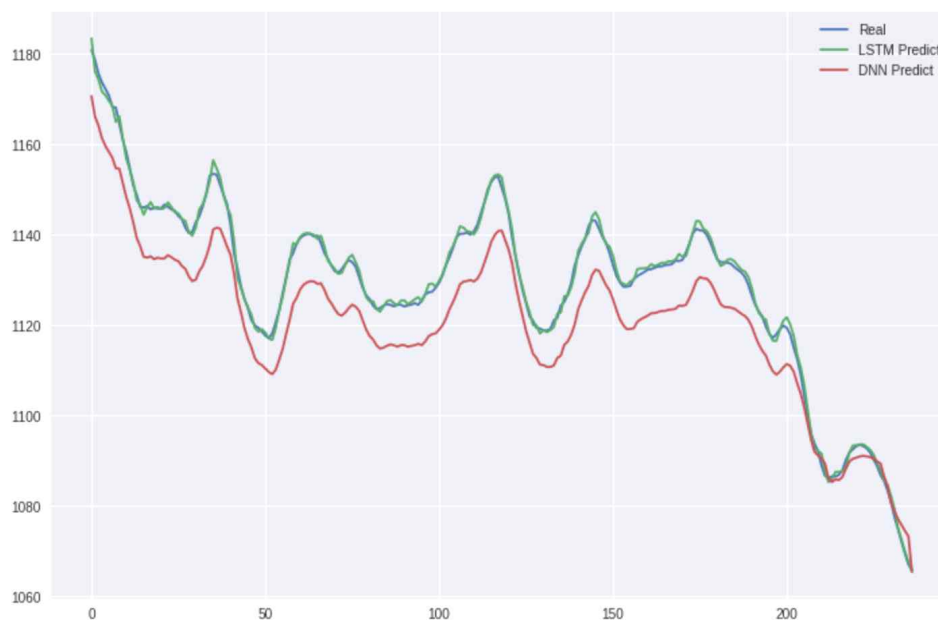
<Figure 15. DA DNN, LSTM Model Predict Plot (2016)>

<Table 13. RMSE, MAPE of DA DNN, LSTM Model (2016)>

	DA LSTM	DA DNN
RMSE	0.017166	0.0603144
MAPE	3.260827	3.2843251

2016년 예측에 사용된 학습데이터는 <Figure 15>와 같이 2011년 1월 3일부터 2015년 12월 30일까지의 데이터이며, 실험 데이터는 2016년 1월 3일부터 2

016년 12월 29일까지의 데이터이다. DA에서도 해당 기간에서 RMSE 수치가 다소 상승하였는데 이는, 해당 기간의 환율 변동성이 다른 기간과 비교하였을 때 규칙적이지 않아 모델의 변동성 제거에 실패하여 예측 성능이 악화된 것으로 해석할 수 있다.



<Figure 16. DA DNN, LSTM Model Predict Plot (2017)>

<Table 14. RMSE, MAPE of DA DNN, LSTM Model (2017)>

	DA LSTM	DA DNN
RMSE	0.00462583	0.03984272
MAPE	1.94256223	2.108413927

2017년 예측에 사용된 학습데이터는 <Figure 16>와 같이 2012년 1월 4일부터 2016년 12월 29일까지의 데이터이며, 실험 데이터는 2017년 1월 2일부터 2017년 12월 28일까지의 데이터이다.



<Figure 17. DA DNN, LSTM Model Predict Plot (2018)>

<Table 15. RMSE, MAPE of DA DNN, LSTM Model (2018)>

	DA LSTM	DA DNN
RMSE	0.00440065	0.02813377
MAPE	2.564800	2.50211283

2018년 예측에 사용된 학습데이터는 <Figure 17>와 같이 2013년 1월 2일부터 2017년 12월 28일까지의 데이터이며, 실험 데이터는 2018년 1월 2일부터 2018년 12월 28일까지의 데이터이다.



<Figure 18. DA DNN, LSTM Model Predict Plot (2019)>

<Table 16. RMSE, MAPE of DA DNN, LSTM Model (2019)>

	DA LSTM	DA DNN
RMSE	0.00433791	0.0599995
MAPE	2.50861046	2.6263366

2019년 예측에 사용된 학습데이터는 <Figure 18>와 같이 2014년 1월 2일부터 2018년 12월 28일까지의 데이터이며, 실험 데이터는 2019년 1월 2일부터 2019년 12월 11일까지의 데이터이다.

다음 <Table 17, 18>에 DA DNN과 DA LSTM 모형의 RMSE 값과 MAPE 값을 예측 기간별로 구분하여 표시하였다.

<Table 17. DA Model Performance (RMSE)>

	전체 기간	15년 예측	16년 예측	17년 예측	18년 예측	19년 예측
DA DNN	0.050027	0.0443292	0.0603144	0.0398427	0.0281337	0.059999
DA LSTM	0.00516	0.0059318	0.017166	0.004625	0.0044006	0.004337

<Table 18. DA Model Performance (MAPE)>

	전체 기간	15년 예측	16년 예측	17년 예측	18년 예측	19년 예측
DA DNN	3.363806	3.592863	3.284325	2.1084139	2.5021128	2.626336
DA LSTM	3.516912	3.83962	3.260827	1.9425622	2.564800	2.508610

DA를 통한 DNN, LSTM 분석에서는 LSTM 예측력이 DNN보다 향상되었음을 알 수 있었다. 특히 LSTM의 예측 성능은 DA의 적용으로 인해 일관되게 향상되었으나, DNN의 경우 예측성능이 악화되거나 크게 개선되지 않는 등 디노이징 인코더의 투입으로 인한 변화가 극적으로 관찰되지는 않았다. 이러한 결과를 통해, LSTM 모형에 투입된 DA는 효과적으로 시계열 변동성으로 인한 노이즈를 줄이는 것으로 해석된다.

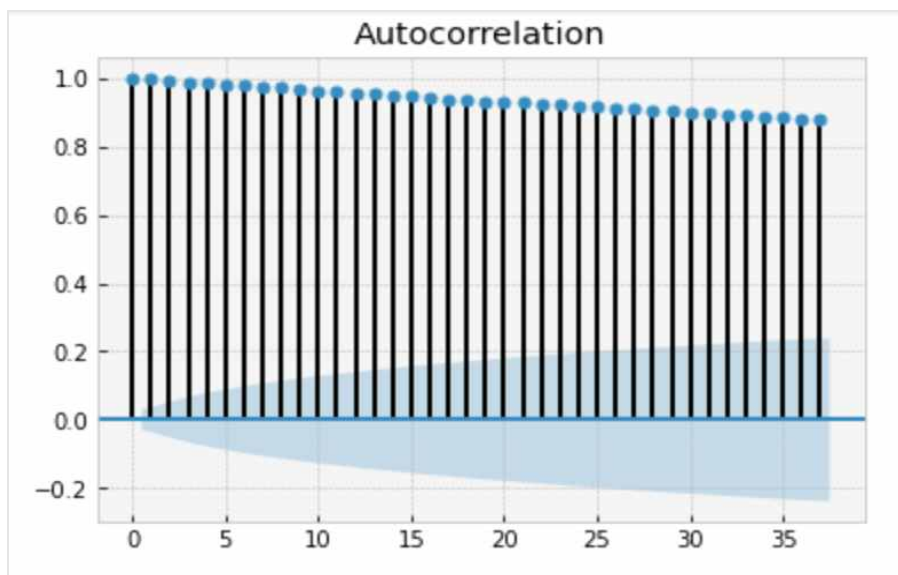
4.4. 추가분석(AR 시계열 분석)

이 연구에서 딥러닝 모형의 예측 결과가 우수함을 입증하기 위해 전통적으로 널리 이용되고 있는 AR(Auto regression Model) 시계열 모형의 예측 결과

를 제시하고자 한다. 시계열 분석에 사용되는 AR 모형은 자기 회귀 모형을 의미한다. 즉 이전의 회귀식에 의한 관측값의 오차량이 이후 관측값에 영향을 주도록 설계된 모형이다. 자기상관계수(Coefficients of AutoCorrelation: AC)와 백색소음(white noise)로 이루어진 회귀방정식을 갖는다.

$$AR: X_t = \phi X_{t-1} + \epsilon_t$$

AR 모형을 사용하기 위해서는 시계열 모형이 AR의 특성을 띠는지 확인해야 한다. AR 특성을 확인하기 위해 ACF(Autocorrelation function)를 활용해 시차에 따른 관측치들 사이에 관련성을 측정할 필요가 있다. 시계열 데이터가 AR 특성을 띠는 경우 ACF는 천천히 감소한다. 반대로 AR 특성이 아닌, MA (Moving Average) 특성을 띠는 경우 AR은 급속도로 감소하게 된다. 아래의 <Figure 19>를 확인하면, 데이터가 AR의 특성을 띠고 있음을 확인할 수 있다.



<Figure 19. Autocorrelation Function Plot>

따라서 AR 모델을 이용하여 파라미터인 환율을 추정하였다. 추정값에 따르면, p-value가 0.1의 유의수준에서 환율의 증가에는 유의하고, 모형의 Constant에는 유효하지 않은 것을 확인할 수 있다. 또한, 모형의 추세를 Constant에서 non-constant로 설정하여 재분석하였다.

```
1 from statsmodels.tsa.arima_model import ARIMA
2
3 model = ARIMA(series[sClose], order=(1,0,0))
4 model_fit = model.fit(trend='c', full_output=True, disp=1)
5 print(model_fit.summary())
```

ARMA Model Results						
Dep. Variable:	Close	No. Observations:	4339			
Model:	ARMA(1, 0)	Log Likelihood	12872.986			
Method:	css-mle	S.D. of innovations	0.012			
Date:	Mon, 19 Oct 2020	AIC	-25739.971			
Time:	13:06:17	BIC	-25720.845			
Sample:	0	HQIC	-25733.219			
	coef	std err	z	P> z	[0.025	0.975]
const	0.3273	0.048	6.845	0.000	0.234	0.421
ar.L1.Close	0.9963	0.001	793.529	0.000	0.994	0.999
Roots						
	Real	Imaginary	Modulus	Frequency		
AR.1	1.0038	+0.0000j	1.0038	0.0000		

<Figure 20. AR1 Table>

<Figure 20>의 결과는 1st order AR, 즉 1기 전의 종속변수가 현시점의 종속변수에 영향을 주고 있는지 보여준다. 유의수준 P가 0.000 즉 매우 유의한 수준에 있고, 1기 전의 종속변수는 0.996의 계수로 유의한 영향을 미치고 있음을 확인할 수 있다.

```
1 from statsmodels.tsa.arima_model import ARIMA
2
3 model = ARIMA(series[sClose], order=(2,0,0))
4 model_fit = model.fit(trend='c', full_output=True, disp=1)
5 print(model_fit.summary())
```

ARMA Model Results						
Dep. Variable:	Close	No. Observations:	4339			
Model:	ARMA(2, 0)	Log Likelihood	12873.922			
Method:	css-mle	S.D. of innovations	0.012			
Date:	Mon, 19 Oct 2020	AIC	-25739.843			
Time:	13:22:51	BIC	-25714.342			
Sample:	0	HQIC	-25730.841			
	coef	std err	z	P> z	[0.025	0.975]
const	0.3269	0.047	6.959	0.000	0.235	0.419
ar.L1.Close	1.0170	0.015	67.018	0.000	0.987	1.047
ar.L2.Close	-0.0208	0.015	-1.368	0.171	-0.051	0.009
Roots						
	Real	Imaginary	Modulus	Frequency		
AR.1	1.0039	+0.0000j	1.0039	0.0000		
AR.2	47.9655	+0.0000j	47.9655	0.0000		

<Figure 21. AR2 Table>

<Figure 20>의 결과는 1st order AR과 2nd order AR, 즉 1기와 2기 전의 종속변수가 현시점의 종속변수에 유의미한 영향을 주고 있는지 보여준다. AR(1)은 유의수준 P가 0.000 즉 매우 유의한 수준에 있고, AR(2)는 0.171로 유의수준 0.05보다 높은 것으로 드러났다. 따라서 2기 전의 종속변수는 현재의 종속변수에 통계적으로 유의한 영향을 미치지 못하는 것으로 드러났다. 또한 AR(1)의 계수는 1.0170으로 영향을 종속변수에 미치고 있음을 확인할 수 있다.

```
1 from statsmodels.tsa.arima_model import ARIMA
2
3 model = ARIMA(series[sClose], order=(3,0,0))
4 model_fit = model.fit(trend='c',full_output=True, disp=1)
5 print(model_fit.summary())
```

ARMA Model Results						
Dep. Variable:	Close	No. Observations:	4339			
Model:	ARMA(3, 0)	Log Likelihood	12874.965			
Method:	css-mle	S.D. of innovations	0.012			
Date:	Mon, 19 Oct 2020	AIC	-25739.930			
Time:	13:28:10	BIC	-25708.053			
Sample:	0	HQIC	-25728.677			
	coef	std err	z	P> z	[0.025	0.975]
const	0.3265	0.046	7.081	0.000	0.236	0.417
ar.L1.Close	1.0165	0.015	66.990	0.000	0.987	1.046
ar.L2.Close	0.0015	0.022	0.071	0.944	-0.041	0.044
ar.L3.Close	-0.0219	0.015	-1.445	0.149	-0.052	0.008
Roots						
	Real	Imaginary	Modulus	Frequency		
AR.1	1.0041	+0.0000j	1.0041	0.0000		
AR.2	6.2894	+0.0000j	6.2894	0.0000		
AR.3	-7.2236	+0.0000j	7.2236	0.5000		

<Figure 22. AR3 Table>

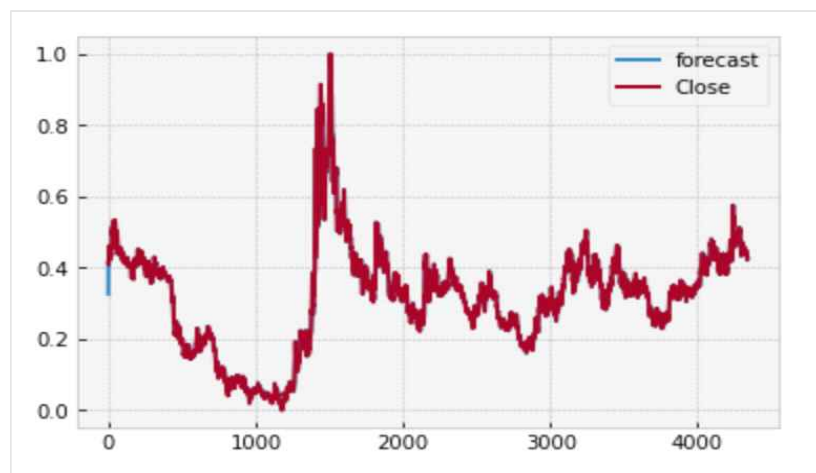
<Figure 20>의 결과는 1st order AR부터 3rd order AR 즉 1기, 2기와 3기 전의 종속변수가 현시점의 종속변수에 주는 영향에 대해 보여준다. AR(1)은 유의수준 P가 0.000 즉 매우 유의한 수준에 있고, AR(2)는 0.994, AR(3)은 0.149로 일반적인 유의수준 0.05보다 높은 것으로 나타났다. 따라서 2기와 3기 전의 종속변수는 현재의 종속변수에 통계적으로 유의한 영향을 미치지 못하는 것으로 해석할 수 있다. 또한 AR(1)의 계수는 1.0165로 영향을 종속변수에 미치고 있음을 확인할 수 있다.

<Table 19. AR Model RMSE, MAPE>

	AR1	AR2	AR3
RMSE	0.2813	0.1685	0.1835
MAPE	7.57	2.34	3.06

<Table 19>는 AR 모형의 회귀 예측 성능을 보여준다. AR 모형은 전반적으로 DNN과 LSTM의 모형과 비교하였을 때, 오차 수준이 높은 것으로 나타났다. 주목할 것은, AR의 lag가 증가한다고 해서 반드시 예측 성능이 좋아지는 않았다. 예측 성능이 가장 우수한 것은 2 시점 전의 종속변수를 자기 회귀에 포함했을 경우로 나타났으며, 이는 회귀 모형의 최적화와 변수의 추가와 소거 등에 의하여 모형이 최적화되는 과정에서 발생한 결과로 해석할 수 있다.

최종적으로 제작한 모형의 예측을 보여주는 그래프는 <Figure 23>과 같다. 딥러닝 기법에 비하여 예측률이 현저히 떨어지는 것을 확인할 수 있다. 따라서 AR 모형의 예측 결과와 비교했을 때 딥러닝 모형의 환율예측 정확도가 매우 우수한 것을 확인할 수 있다.



<Figure 23. AR2 Model Predict Plot>

제5장. 결론

5.1. 결론 및 시사점

인공지능(AI) 기법을 활용해 다양한 금융 시계열 변수들에 관한 예측 연구들이 최근 활발히 이루어지고 있다. 인공지능 기법을 활용한 연구가 활발해지는 것은 계량 모형들의 한계가 드러나면서 기업의 재무 담당자들은 리스크 관리를 위해 더욱 정확한 모형이 필요하게 되었으며, 과거에 이론적으로 제시되었으나 당시 컴퓨팅 파워의 부족으로 사용하기 어려웠던 인공신경망 기법들이 재조명받기 시작하였다.

이 논문에서는 인공신경망 기법 중 DNN과 LSTM을 이용해 환율 증가에 관한 예측 연구를 수행하였다. 전체 기간에서 DNN 모형보다 LSTM 모형이 더 좋은 예측성능을 보였으며, 특히 DA(De noising Autoencoder)를 적용하여 슬라이딩 윈도우 기법으로 구간을 구분하여 연구를 수행하였을 때 LSTM이 더욱 우수한 예측 성능을 보였다. LSTM의 분석 결과는 DA의 적용에 따라 오차가 감소하는 결과를 보였지만, DNN은 분석 기간에 따라 오히려 오차가 증가하는 경우도 존재하였다. 이러한 결과를 바탕으로, LSTM에 적용된 DA는 효과적으로 시계열 노이즈를 통제하는 것으로 해석할 수 있다. 반면 DNN은 LSTM에 비하여 시계열 학습 능력이 떨어지며, 시계열 노이즈의 제거가 반드시 모델의 예측성능을 높이고 있지는 않은 것으로 해석된다. 또한, 고전 시계열 예측 모형인 AR 모형과 비교하였을 때 LSTM과 DNN이 더욱 우수한 예측 성능을 보였다. LSTM이 시계열 분석에서 강점을 가진 것으로 알려져 있듯이, 본 연구 결과에서도 전반적으로 좋은 예측성을 보였다. 환율 데이터의 예측 성능을 높이기 위해서는 환율의 시가, 저가, 고가,

종가, 거래량 등의 기술적 지표 이외의 추가적으로 시장의 변동성을 반영하는 지표가 필요할 것이다.

하지만, 본 연구에는 국내의 환율 예측 연구 중 처음으로 LSTM 모델을 적용하였다는 의의와 환율 예측연구의 정확성을 높이기 위해서는 환율의 변동성을 반영하고 있는 다양한 정량적, 정성적 변수들이 고려되어야 함을 입증하였다는 것에 큰 의의가 있다.

5.2. 한계 및 향후 연구

국내의 환율 제도는 1997년 외환위기(IMF) 이후 자유 변동환율제도로 변경되었다. 이에 따라 누적된 환율 데이터의 절대적인 양이 부족한 상황이다. 또한 외환위기(IMF) 10년 후인 2008년에는 서브프라임 모기지(Subprime mortgage) 사태로 또다시 외환시장에 커다란 변화의 바람이 불었다. 이 논문에서 불확실성과 잡음을 최소화하기 위해 2010년 이후의 일별 데이터를 표본으로 활용하였기 때문에 표본 수가 부족하다는 한계가 있다. 그리고 환율의 기술적 지표만을 활용하여 모델을 설계했기 때문에 다소 단순한 구조의 모델을 사용하였다는 한계가 있다.

따라서 향후 연구에서는 환율 예측에 정량적, 정성적 변수들을 추가로 고려하여 연구를 진행해야 할 것이다. 또한 변수를 확보한 후 현재의 단순한 구조의 모델을 확장시켜 양방향 LSTM, Conv-LSTM 등 최신 인공지능 기법을 적용하여 연구를 수행할 수 있을 것으로 생각한다.

참 고 문 헌

- [1] Ahn. J. J., Byun, H. W., Oh, K. J. and Kim, T. Y. (2012). *Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting*. Expert Systems with Applications, 39, 8369-8379.
- [2] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Saverese, S. (2016). *Social lstm: Human trajectory prediction in crowded spaces*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 961-971).
- [3] C. Lee, C. Chang and F. Hwang(2019), "*Currency Exchange Rate Prediction with Long Short-Term Memory Networks Based on Attention and News Sentiment Analysis*," 2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), Kaohsiung, Taiwan, pp. 1-6.
- [4] Chen, K., Zhou, Y., & Dai, F. (2015, October). *A LSTM-based method for stock returns prediction: A case study of China stock market*. In 2015 IEEE international conference on big data (big data) (pp. 2823-2824). IEEE.
- [5] Cocianu, C. L., & Avramescu, M. Ș. (2020). The Use of LSTM Neural Networks to Implement the NARX Model. A Case Study of EUR-US D Exchange Rates. *Informatica Economica*, 24(1), 5-14.
- [6] Clarida, R. H., & Taylor, M. P. (1997). The term structure of forward exchange premiums and the forecastability of spot exchange rates: correcting the errors. *Review of Economics and Statistics*, 79(3), 353-361.

- [7] Cheong, D. H. and Oh, K. J.. (2014). Using cluster analysis and genetic algorithm to develop portfolio investment strategy based on investor information. *Journal of Korean Data & Information Science Society*, 25:1, 107–117.
- [8] Diebold, F. X., & Nason, J. A. (1990). *Nonparametric exchange rate prediction?*. *Journal of international Economics*, 28(3–4), 315–332.
- [9] Dixon, M., Klabjan, D., & Bang, J. H. (2017). *Classification-based financial markets prediction using deep neural networks*. *Algorithmic Finance*, 6(3–4), 67–77.
- [10] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.
- [11] Fischer, T., & Krauss, C. (2018). *Deep learning with long short-term memory networks for financial market predictions*. *European Journal of Operational Research*, 270(2), 654–669.
- [12] Fu, R., Zhang, Z., & Li, L. (2016, November). *Using LSTM and GRU neural network methods for traffic flow prediction*. In 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC) (pp. 324–328). IEEE.
- [13] Galeshchuk, S. (2016). *Neural networks performance in exchange rate prediction*. *Neurocomputing*, 172, 446–452.
- [14] Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.
- [15] Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). *Learning to forget: Continual prediction with LSTM*.

- [16] Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). *Learning precise timing with LSTM recurrent networks*. Journal of machine learning research, 3(Aug), 115–143.
- [17] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). *LSTM: A search space odyssey*. IEEE transactions on neural networks and learning systems, 28(10), 2222–2232.
- [18] Huang, Z., Xu, W., & Yu, K. (2015). *Bidirectional LSTM-CRF models for sequence tagging*. arXiv preprint arXiv:1508.01991.
- [19] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527–1554.
- [20] Ince, H., & Trafalis, T. B. (2006). *A hybrid model for exchange rate prediction*. Decision Support Systems, 42(2), 1054–1062.
- [21] Jeong, G., & Kim, H. Y. (2019). *Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning*. Expert Systems with Applications, 117, 125–138.
- [22] Jeong, K. H. (2017). *Quantile causality from dollar exchange rate to international oil price*. Journal of the Korean Data & Information Science Society, 28, 361–369.
- [23] Kim, S. H. (2000). *Establishment of Optimal Artificial Neural Network Model and Exchange Rate Prediction Performance Analysis*. Journal of Money and Finance, 14, 57–85.
- [24] Kim, T. Y., Oh, K. J., Sohn, I. and Hwang, C. (2004). *Usefulness of artificial neural networks for early warning system of economic crisis*. Expert Systems with Applications, 26, 583–590.

- [25] Kim, E. C., & Oh, K. J. (2019). Asset allocation strategy using hidden Markov model and genetic algorithm. *The Korean Data & Information Science Society*, 30(1), 33-44.
- [26] Lee, J. Y. and Kim, H. J. (2014). *Identification of major risk factors associated with respiratory diseases by data mining*. Journal of the Korean Data & Information Science Society, 25, 373-384.
- [27] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- [28] Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. *In Text summarization branches out* (pp. 74-81).
- [29] Meese, R. A. and Rogoff, K. (1983). *Empirical exchange rate models of the seventies: Do they fit out of sample?*. Journal of International Economics, 14, 3-24.
- [30] Refenes, A. N., Azema-Barac, M., Chen, L., & Karoussos, S. A. (1993). *Currency exchange rate prediction and neural network design strategies*. Neural Computing & Applications, 1(1), 46-58.
- [31] Park, S. and Park, C., "Time-varying Cointegration Models and Exchange Rate Predictability in Korea," *KDI Journal of Economic Policy*, 37, 2015, pp.1-20.
- [32] Shin, Y. G. (2009). *Study on the causality between call rate and exchange rate under global economic crisis*. Journal of the Korean Data & Information Science
- [33] Varian, H. R. (2014). *Intermediate microeconomics with calculus: a modern approach*. WW Norton & Company.

- [34] Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008, July). Extracting and composing robust features with denoising autoencoders. *In Proceedings of the 25th international conference on Machine learning* (pp. 1096-1103).
- [35] W. Cao, W. Zhu, W. Wang, Y. Demazeau and C. Zhang(2020), "A Deep Coupled LSTM Approach for USD/CNY Exchange Rate Forecasting," in *IEEE Intelligent Systems*.
- [36] Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1991, November). *Generalization by weight-elimination applied to currency exchange rate prediction*. In *[Proceedings] 1991 IEEE International Joint Conference on Neural Networks* (pp. 2374-2379). IEEE.
- [37] Zhang, G. P., & Berardi, V. L. (2001). *Time series forecasting with neural network ensembles: an application for exchange rate prediction*. *Journal of the operational research society*, 52(6), 652-664.
- [38] 강민영. (2017). **CNN기반 원/달러 환율예측 모형 연구**. 석사학위논문, 서강대학교 정보통신대학원.
- [39] 김영철, 이현준, 김지우, 안재준. (2018). **스왑포인트 결정요소를 이용한 머신러닝 기반의 원/달러 환율 예측 모형에 관한 연구**. 한국데이터정보과학회지, 29(1), 203-216. ience Society, 20, 655-660.
- [40] 김재현. (2001). **인공신경망 모형과 ARIMA 모형의 원/달러 환율예측 성과 비교연구**. 서강대학교 대학원, 51-60.
- [41] 박혜원. (2019). **시계열모형의 예측에서 LSTM 성능에 대한 연구**. 석사학위논문, 계명대학교 대학원.
- [42] 신동하, 최광호, & 김창복. (2017). **RNN 과 LSTM 을 이용한 주가 예측을 향상을 위한 딥러닝 모델**. 한국정보기술학회논문지, 15(10), 9-16.

- [43] 이낙영, 오경주 (2019) 디노이징 필터와 LSTM을 활용한 KOSPI200 선물지수 예측. 한국데이터정보과학회지.

Abstract

Proposal of Won/Dollar Exchange Rate Prediction Model Using Deep Learning Technique - Focused on DNN and LSTM -

Byun, Hyun jung

Major in Investment Information Engineering

The Graduate School of Engineering

Yonsei University

Supervised by Prof. ____

Exchange rates are a very important variable that has a profound effect on trade between countries. In particular, foreign exchange risk has a significant impact on the valuation of financial and real assets. Predicting exchange risks is a very important issue due to the high dependence on exports in Korea. However, the limitations of the existing financial time series model, which was widely used for forecasting in the financial and macro economic sectors, have been revealed, and under these circumstances, research using the deep learning model, which shows strength in analyzing time series, is becoming more active. However, there is a lack of research on currency forecasts in Korea. This study presents a model of exchange rate prediction using DNN and LSTM, which has never been attempted in existing domestic studies.

The exchange rate prediction models presented in this study used the quoted exchange rate price, high price, low price, closing price and trading volume as input variables, and the DNN and LSTM models were used as a way to deal with vast amounts of data. The empirical analysis of the two models was conducted using the sliding window technique and the data for 10 years from 2010. In order to further improve the performance of the predictive model, the purpose of this study was to utilize the Denoising Autoencoders. After eliminating noise from the data using DA, we expect to improve the time series prediction performance of DNN and LSTM.

In this paper, the learning period was set at seven years for the entire data to optimize the model. The exchange rate forecasts were carried out using the two models subsequently designed, and RMSE and MAPE were used as indicators to evaluate the predicted performance of both DNN and LSTM models. In addition to the deep learning model, the addition of the AR model as a benchmark model demonstrated the excellence of the model presented in this study. Forecasts show LSTM's performance is better than that of DNN. This study is the first of its kind to utilize LSTM among domestic exchange rate prediction studies, and is meaningful in providing clues to future exchange rate prediction research.

Keywords: exchange rate prediction, sliding window, deep neural network, short-term and long-term memory circulation neural network, denoising autoencoder, Autoregression model