텍스트 감정 분석과 딥러닝 기반 시계열 예측 모델을 활용한 USD/KRW 환율 예측 연구

Jaeheung Parka, Kyungwon Kima,*∗*

*aSchool of International Trade and Business, Incheon National University, Incheon, 22012, Republic of Korea*

# Abstract

*Background:* Traditional exchange rate forecasting models often rely solely on macroeconomic indicators, overlooking investor sentiment and market psychology, which play a critical role in short-term price movements. With the growing availability of unstructured financial text data, sentiment analysis has emerged as a promising approach to improve financial forecasting models.

*Objective:* This study aims to investigate whether integrating sentiment scores extracted from news articles and investor comments into deep learning-based time series models can enhance the prediction accuracy of the USD/KRW exchange rate.

*Participants and Methods:* We collected news articles related to exchange rates from Naver Finance and USD/KRW comments from Investing.com, covering the period from January 2020 to December 2024. Sentiment scores were derived using a fine-tuned KR-FinBERT-SC model, trained on approximately 300 manually labeled samples. Sentiment data were categorized into positive, negative, and neutral labels (3-class) or positive and negative labels (2-class), and aggregated daily. The final dataset consisted of macroeconomic indicators combined with or without sentiment scores. We employed four deep learning models—LSTM, GRU, CNN-LSTM, and CNN-GRU—trained on 80% of the data and evaluated on the remaining 20%. All numerical features were scaled using MinMax normalization, and model performance was assessed using MSE and MAPE.

*Results:* Models that incorporated sentiment variables consistently outperformed those using only economic indicators. The GRU model showed the most stable and accurate performance across different input combinations. Notably, models using two-class sentiment scores from news or high-confidence scores (≥0.9) from comments achieved the lowest prediction errors, with the GRU model reaching an MSE as low as 65.01 and a MAPE of 0.57. In contrast, CNN-based models showed comparatively weaker performance, indicating limitations in capturing long-term dependencies in time series data.

*Conclusions:* The results demonstrate that market sentiment derived from unstructured text data can significantly improve exchange rate forecasting. By combining sentiment analysis with time series deep learning models, particularly GRU, it is possible to capture both economic fundamentals and psychological factors influencing the market. This study provides empirical evidence for the effectiveness of emotion-aware models in financial forecasting and suggests future research directions involving diverse data sources and advanced architectures such as Transformers.

*Keywords:* Exchange Rate, Sentiment Analysis, Deep Learning

*∗*Corresponding author: thekimk.kr@gmail.com

*ORCID(s):* 0000-0001-6530-8426 (K. Kim)

# 서론

환율(Exchange rate)이란 외국 통화 한 단위를 받기 위해 자국 통화를 몇 단위 지불해야 하는가를 나타내는 것으로 자국 통화와 외국 통화간의 교환비율을 의미하며 두 나라 통화의 상대적 가치를 말한다(한국은행). 오늘날 거의 모든 경제 활동은 국제 거래와 밀접하게 연계되어 있기 때문에, 환율은 일상생활은 물론 기업 경영과 국가 경제 운영에 이르기까지 필수적인 정보로 기능한다(국제경제론 김인준, 이영섭).

특히 USD/KRW 환율은 한국의 대외경제에 큰 역할을 미치는 요소로, 수출입 가격, 외국인 투자, 물가 수준, 그리고 중앙은행의 통화정책 등 다양한 거시경제 변수와 밀접하게 연관되어 있다. 이에 따라 환율의 향후 변화를 예측하는 것은 국가 경제의 안정성과 기업의 재무 전략 수립에 있어 핵심적인 과제로 인식되고 있다.

그러나 환율은 경제 펀더멘털뿐만 아니라 지정학적 리스크, 글로벌 자본 흐름, 시장 심리 등 복합적 요인의 영향을 받으며, 이러한 불확실성으로 인해 종종 ‘랜덤워크(Random Walk)’로 설명된다. 랜덤워크란 매우 술에 취한 사람이 어디로 걸어갈지 예측할 수 없는 것처럼, 환율 또한 다음 시점의 방향을 정확히 예측하기 어렵다는 비유적 표현으로 사용된다. (https://www.hankyung.com/article/2023102265501)

환율은 외환시장의 수요와 공급에 따라 결정되고, 각국의 물가수준, 생산성 등 경제여건의 변화는 장기적으로 통화의 가치에 영향을 미친다고 한다. 중기적 관점에서 보면 환율에 영향을 미치는 요인으로 대외거래, 거시경제정책 등을 들 수 있으며, 단기적으로 환율은 외환시장 참가자들의 기대나 주변국의 환율 변동, 각종 뉴스 등에 따라 영향을 받는다고 한다.

특히, 시장참가자들의 환율에 대한 기대가 변하게 되면 자기실현적(Self-fulfilling)인 거래에 의해 실제 환율의 변동이 초래된다고 한다. 예를 들어 대부분의 시장참가자가 환율상승을 예상할 경우 환율이 오르기 전에 미리 외환을 매입하면 이익을 볼 수 있으므로 외환에 대한 수요가 증가하게 되어 실제 환율이 상승하게 된다.

뿐만 아니라, 각종 뉴스도 시장참가자들의 기대변화를 통해 단기 환율변동에 영향을 미치게 된다고 한다. 일례로 2010년 5월 천안함 침몰조사 결과가 발표되고 지정학적 위험이 부각되자 원/달러 환율이 일시적으로 큰 폭 상승하였다(한국은행).

(https://www.bok.or.kr/portal/main/contents.do?menuNo=200407)

그렇다면, 시장 참여자들의 감정과 뉴스의 감정을 반영한 감정 분석(Sentiment Analysis)을 통해 환율을 보다 정교하게 예측할 수 있지 않을까? 본 연구는 이러한 질문에서 출발하여, 감정 분석 결과를 딥러닝 기반 시계열 예측 모델에 결합함으로써 USD/KRW 환율 예측 정확도 향상 가능성을 탐색하고자 한다.

1. 선행 연구

2.1 감정분석을 통한 주가 예측 연구

정가연 et al. (2022)는 뉴스 데이터를 기반으로 감정 분석 모델과 딥러닝 기법을 결합하여 코스피 지수의 종가를 예측하였다. 이들은 독립 변수로 코스피 고가, 다우존스 지수 거래량, S&P500 지수 저가 및 거래량, 유가 시가 및 거래량, 금값 저가, 원/엔 환율 저가, 한국 기준금리 등 주요 거시경제 지표와 함께, 코스피 검색 빈도 및 감정 점수 등 시장 심리를 반영하는 변수들을 활용하였다. 특히 KLUE-BERT 기반의 감정 분석 결과를 시계열 입력 변수로 전환하여 LSTM, GRU, CNN-LSTM, CNN-GRU 모델에 투입하였으며, 뉴스 본문 요약을 통해 산출한 감정 점수를 CNN-GRU 모델에 적용했을 때 가장 우수한 성능을 보였다.

양혁진 et al.(2021)는 비정형 뉴스 데이터와 정형 데이터를 통합하여 주가 지수를 예측하였다. 이들은 네이버 뉴스와 증권사 리포트를 바탕으로 긍·부정 감정 점수를 추출하고, GRU 및 LSTM 기반의 딥러닝 모델에 적용하였다. 그 결과, 감정 분석 결과를 포함한 모델이 포함하지 않은 모델에 비해 실제 주가 예측에서 더 뛰어난 성능을 나타냈다.

김지환 et al.(2023)은 The New York Times의 기사를 요약한 후 FinBERT 모델을 통해 감정 분석을 수행하였으며, 도출된 감정 점수를 기반으로 LSTM 모델을 통해 S&P500 지수를 예측하였다. 실험 결과, 감정 점수를 반영한 모델이 더 높은 예측 정확도를 보였다.

2.2 딥러닝을 통한 환율 예측 연구

한편, Wei Cao et al.(2020)은 USD/CNY 환율 예측을 위해 DC-LSTM 모델을 제안하였다. 이들은 유가, 금값, 상하이 종합지수, 다우존스 지수, 중국의 통화공급지표(M1, M2), 소비자물가지수, 생산자물가지수, 산업생산지수, 기준금리, 인플레이션율, 무역수지 및 국제수지 지수, 정책 불확실성 지수 등 다양한 경제 지표를 독립 변수로 활용하였다. 실험 결과, DC-LSTM 모델이 LSTM, CNN, SVR, ARIMA 모델보다 우수한 예측 성능을 보였다.

2.3 선행 연구 시사점

이러한 선행연구들은 두 가지 중요한 시사점을 제공한다. 첫째, 정가연 et al.(2022), 양혁진 et al.(2021), 김지환 et al.(2023)의 연구는 감정 분석 결과를 예측 모델에 추가함으로써 시계열 예측의 성능이 향상될 수 있음을 보여주었다. 특히, 텍스트 데이터를 기반으로 추출된 감정 점수를 딥러닝 모델에 추가함으로써 보다 우수한 예측이 가능함을 입증하였다.

둘째, Wei Cao 외(2020)의 연구는 유가, 금값, 주요 주가지수, 통화 공급량 등 다양한 경제 지표들을 활용하여 환율 예측이 가능하다는 점을 실증적으로 보여주었다.

감정 분석을 활용한 비정형 데이터의 통합은 예측 모델의 성능을 높이는 데 효과적이며, 동시에 경제 펀더멘털 변수들은 여전히 환율 변동의 중요한 설명 변수로 기능하고 있음을 알 수 있다. 이에 따라 본 연구는 이러한 두 가지 접근을 결합하여, 감정 분석과 주요 경제 지표를 통합한 딥러닝 모델을 통해 환율 예측의 정확도를 높이고자 한다.

# 연구 방법론

* 1. *Participants and Data Preprocessing*

본 연구에서는 USD/KRW 환율의 종가를 종속 변수로 설정하고, 이를 예측하기 위한 독립 변수로 다음과 같은 거시경제 지표 및 시장 심리 지표를 활용하였다.

거시경제 지표에는 USD/KRW 시가, KOSPI 종가 및 거래량, KOSDAQ 종가 및 거래량, 국제 유가 종가, 금 가격 종가, 다우존스 지수 종가 및 거래량, S&P500 종가, 한국의 통화공급량(M1, M2), 한국 및 미국의 정책금리, 소비자물가지수(CPI), 생산자물가지수(PPI), 경상수지 등이 포함된다.

또한, 시장 참여자의 심리를 반영하기 위해 뉴스 및 댓글 기반 감정 점수를 추가적인 독립 변수로 포함하였다.

데이터는 2020년 1월 1일부터 2024년 12월 31일까지의 기간을 대상으로 수집하였으며, 일 단위로 통합하였다. 이때, 금융시장이 실제로 개장한 영업일 기준 데이터만을 사용하여 분석의 정확성과 현실 반영성을 높였다.

표 1. DB

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **데이터출처** | **데이터베이스(DB)명** | **기간** | **단위** | **추출변수** | **종속변수활용** |
| **인베스팅닷컴** | **USD/KRW** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가 | USD/KRW 종가, USD/KRW 시가, |
| **KOSPI** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가, 거래량 | KOSPI 종가, KOSPI 거래량 |
| **KOSDAQ** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가, 거래량 | KOSDAQ 종가, KOSDAQ 거래량 |
| **WTI** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가, 거래량 | WTI 종가 |
| **XAU** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가, 거래량 | XAU 종가 |
| **다우존스** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가, 거래량 | 다우존스 종가, 다우존스 거래량 |
| **S&P 500** | 2020.01-2024.12 | 일 | 거래일, 종가, 시가, 거래량 | S&P500 종가 |
| **ECOS, FRED** | **정책금리** | 2020.01-2024.12 | 월 | 날짜, 한국 정책 금리, 미국 정책 금리 | 한국 정책 금리, 미국 정책 금리 |
| **통계청** | **소비자 물가지수** | 2020.01-2024.12 | 월 | 날짜, 소비자 물가지수 | 소비자 물가지수 |
| **한국은행** | **생산자 물가지수** | 2020.01-2024.12 | 월 | 날짜, 생산자 물가지수 | 생산자 물가지수 |
| **한국은행** | **경상수지** | 2020.01-2024.12 | 월 | 날짜, 경상수지 | 경상수지 |
| **한국은행** | **통화 공급 통화량** | 2020.01-2024.12 | 월 | 날짜, M1, M2 | M1, M2 |

감정 분석을 위한 텍스트로는 네이버 금융 뉴스 중 환율 카테고리에 해당하는 기사를 2020년 1월 1일부터 2024년 12월 31일까지 크롤링을 진행하였다. 또한 인베스팅 닷컴의 USD/KRW 댓글을 2020년 1월 1일부터 2024년 12월 31일까지 크롤링을 진행하였다.

3.2 뉴스 데이터 수집 및 전처리

본 연구에서는 감정 분석을 위한 텍스트 데이터로 네이버 금융의 ‘환율’ 카테고리에 속하는 기사를 2020년 1월 1일부터 2024년 12월 31일까지 수집하였다. 수집된 기사에 대해 다음과 같은 전처리 과정을 수행하였다. 첫째, 정규표현식을 활용하여 특수문자 및 불필요한 패턴을 제거하고, 본문이 비어 있거나 USD/KRW 환율과 관련 없는 기사(예: 북한 관련 기사)를 제거하였다. 둘째, 기사 본문을 KoBART-Summarization 모델을 활용해 요약하여 감정 분석에 적합한 형태로 변환하였다.

3.3 뉴스 감정 분석 및 모델 파인튜닝

감정 분석 모델로는 금융 도메인에 특화된 KR-FinBERT-SC 모델을 사용하였다. 다만 환율 관련 텍스트에서 성능이 떨어지는 문제가 있어, 직접 라벨링한 208개(3분류, 2분류 둘 다 208개)의 뉴스 문장을 바탕으로 파인튜닝을 진행하였다.  
라벨링 기준은 다음 표와 같다.

|  |  |
| --- | --- |
| **감정** | **분류 기준** |
| 긍정 | 시장 안정, 환율 하락, 외국인 투자 확대, 호재 반응, 강세, 회복, 순매수 등 |
| 부정 | 시장 불안, 환율 급등, 지정학적 리스크, 악재 반응, 약세, 순매도 등 |
| 중립 | 방향성 없는 기술적 설명, 단순 팩트 전달, 보합 등 |

파인튜닝은 긍·부정·중립 3분류 모델과 긍·부정 2분류 모델로 각각 수행하였으며, 최종적으로 요약된 뉴스 본문에 대해 감정 분석을 진행하였다. 감정 점수는 긍정: +1, 중립: 0, 부정: -1로 매핑하여 일자별로 평균값을 계산하였다. 이후 Z-score 정규화를 통해 감정 점수를 표준화하였다. 면밀한 분석을 위해 감정 score가 0.95 이상, 0.90 이상, 전체 집단의 세 가지 기준으로 나누어 별도로 분석을 진행하였다.

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.

**3.4 댓글 데이터 수집 및 감정 분석**

투자자 심리를 직접적으로 반영할 수 있는 데이터로서, Investing.com의 USD/KRW 관련 댓글을 2020년 1월 1일부터 2024년 12월 31일까지 수집하였다. 원본 KR-FinBERT-SC 모델은 대부분의 댓글을 중립으로 분류하는 경향이 있어, 339개의 댓글을 ‘공포’ 또는 ‘욕심’으로 직접 라벨링한 후 모델을 파인튜닝하였다. 감정 점수는 공포: -1, 욕심: +1로 매핑하고, 일자별 평균을 계산하여 감정 점수를 도출하였다. 또한 해당 일자의 댓글 수 정보도 함께 수집하였다. 감정 score에 따라 0.90 이상 집단과 전체 집단으로 구분해 분석을 진행하였다. 0.95 이상 집단은 데이터 수가 부족하여 분석에서 제외하였다. 이후 Z-score 정규화를 통해 감정 점수를 표준화하였다.

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.

* 1. *Deep Learning Algorithm: LSTM, GRU*

*\** ***금융 특화 감정분석 모델과 딥러닝 시계열 예측 모델을 활용한 코스피 지수 예측 Copy***

최근에는 시계열 데이터를 효율적으로 예측하는 신경망 구조 기반의 딥러닝 모델 주가 예측 연구들이 수행되고 있다. 순환신경망(Recurrent Neural Network, RNN)이란 일반 신경망 구조에 시계열 데이터 개념이 추가된 구조로 은닉층에 존재하는 과거 정보를 활용한다(Elman, 1990). LSTM(Long-Short Term Memory)은 RNN의 한 종류로 시퀀스 데이터를 처리하고 장기 의존성(Long-Term Dependencies)을 효과적으로 학습할 수 있으며, memory cell, input gate,forget gate, output gate로 구성되어 있다. LSTM은 memory cell을 통해서 각 시간 단계에서 내부 상태를 유지하고 필요한 정보를 유지함으로써 장기 의존성 문제를 해결했다(Hochreiter and Schmidhuber, 1997). GRU(Gated Recurrent Unit)는 RNN의 한 종류로 update gate와 reset gate로 구성된다.

LSTM보다 더 간단한 구조를 가지고 있으며, 계산 비용이 적게 들지만 비슷한 성능을 보이며 효과적으로 작동한다(Chung et al.,2014). LSTM 기반의 시계열 예측 모델을 학습할 때 합성곱 연산을 수행하는 층(convolution layer)을 결합한 연구가 있으며 단순 LSTM을 활용했을 때보다 높은 예측 정확도를 가진다는 결과가 나왔다(Hwang and Shin, 2020). LSTM 및 GRU는 주기성이나 특정 시간 스케일의 특징을 감지하는 데 어려움이 존재한다. LSTM에 데이터를 입력하기 전 합성곱 연산을 수행하면 주요한 시간적 패턴을 학습할 수 있다. 이를 통해 보다 유용한 표현을 제공할 수 있으며 시계열 데이터에 포함된 잡음이나 이상치를 줄일 수 있다. 신경망 모델을 사용하여 금융 시계열 예측 모델링을 처음 시도한 연구에서는 다양한 딥러닝 모델을 적용해서 IBM의 주식 가격 변동성에서 패턴을 찾아내는 신경망 모델링을 진행하였고 효율적 시장 가설에 대한 증거를 확립하였다(White,1988). 또한 주가 예측 모델로서 LSTM을 적용할 때 성능향상을 위해 고려해야 할 다양한 파라미터 설정과 함수들에 대한 적절한 조합 방법을 제안한 연구에서는 주가 예측을 위한 LSTM 적용 시 최적의 모델링 방법을 실증적인 형태로 제안하였다(Jung and Kim, 2020). CNN 기반 프레임워크를 활용해서주가를 예측하기 위한 특징을 추출한 연구에서는 S&P 500, NASDAQ, Dow Jones Index(DJI) 시장 지수의 다음 날 이동 방향을 예측하였고 기존 알고리즘에 비해 예측 성능이 크게 증가하였다(Hoseinzade and Haratizadeh, 2019). 이러한 방법론들은 기존 통계적 모델을 활용한 모델보다 복잡한 모델을 사용할 수 있고, 모델링을 하기 위한 가정이 상대적으로 덜 필요하다는 장점이 있다. 그러나, 주가 예측시 시장 참여자들의 반응들을 전혀 고려하지 못하고 과거 주가 데이터만 활용한다는 한계가 존재한다.



그림 4. 대표적인 딥러닝 알고리즘인 MLP의 구조와 데이터 학습을 통한 수많은 가중치 추정 과정

* 1. *Hybrid Deep Learning Algorithm: CNN-LSTM, CNN-GRU*

딥러닝 기반 시계열 예측 모델은 데이터의 시간적 패턴뿐만 아니라, 입력 변수 간의 복잡한 구조적 관계를 효과적으로 반영할 수 있어 최근 다양한 연구에서 활발히 활용되고 있다. 특히 CNN-LSTM과 CNN-GRU는 Convolutional Neural Network(CNN)의 지역적 특징 추출 능력과 Recurrent Neural Network(RNN)의 시계열 정보 처리 능력을 결합한 하이브리드 모델로, 시계열 데이터에 포함된 다차원적인 특징을 보다 정교하게 학습할 수 있는 구조를 갖는다.

CNN-LSTM 모델은 입력 시계열 데이터에 대해 먼저 1차원 합성곱 층(1D Convolution Layer)을 통해 국소적이고 유의미한 특징(feature)을 추출한 후, 이를 LSTM 레이어에 전달하여 시간 흐름에 따른 장기적인 의존 관계를 학습한다. 이 과정에서 CNN은 변수 간 상호작용이나 패턴의 공간적 특징을 요약하고, LSTM은 이를 바탕으로 시계열적 특성을 학습함으로써 예측 성능을 강화한다 (Shi et al., 2015).

CNN-GRU 모델 또한 유사한 구조를 가지며, CNN으로부터 추출된 feature map을 GRU 레이어에 전달하여 예측을 수행한다. GRU는 LSTM보다 구조가 간결하면서도 학습 속도가 빠르다는 장점이 있으며, 적은 데이터에서도 안정적인 성능을 보인다.

이러한 CNN-RNN 계열의 하이브리드 모델은 다음과 같은 측면에서 환율 예측에 효과적이다. CNN은 뉴스 및 댓글로부터 생성된 감정 점수를 비롯한 다양한 입력 변수 간의 상관 구조를 파악하여 요약된 특징을 생성한다. LSTM 및 GRU는 해당 특징들의 시간적 변화 흐름을 분석하여 미래 환율 값을 예측한다.

따라서 본 연구에서는 감정 분석 기반의 시계열 데이터를 CNN-LSTM 및 CNN-GRU 모델에 입력함으로써, 시장 심리의 국소적 변화와 장기 흐름을 동시에 반영할 수 있는 예측 시스템을 구축하였다. 특히 뉴스 요약 및 댓글 감정 점수와 같은 비정형 요소들이 고차원의 특성으로 변환되어 CNN에서 효과적으로 추출되고, 이로 인해 환율 예측의 정밀도가 향상되는 효과를 기대할 수 있다.

* 1. *Evaluation Metrics*

본 연구에서는 USD/KRW 환율 예측 모델의 성능을 정량적으로 평가하기 위해 MSE(Mean Squared Error)와 MAPE(Mean Absolute Percentage Error)의 두 가지 회귀 성능 지표를 사용하였다. 이들 지표는 예측값과 실제값 간의 차이를 기반으로 하며, 값이 낮을수록 예측 정확도가 높다는 것을 의미한다. **MSE**는 예측 오차의 제곱 평균으로, 큰 오차에 더 큰 패널티를 부여함으로써 모델의 전반적인 예측 정밀도를 측정한다. **MAPE**는 실제값에 대한 예측 오차의 비율을 평균한 값으로, 예측 오차를 상대적인 퍼센트로 표현하여 해석이 직관적이다.

모든 metrics는 예측 성능이 좋을수록 낮은 수치들이 나오도록 실제값()과 예측값()의 차이로 구성되어 있다. 따라서 2개의 metrics 모두 낮은 수치가 나오는 경우 환율을 잘 예측하는 모델이라 볼 수 있다.

# Results

본 연구에서는 감정 분석 관련 변수를 제외한 전통적인 경제 지표만을 활용한 예측 모델과, 감정 분석 변수를 포함한 모델의 예측 성능을 비교하고자 한다. 이를 통해 감정 분석이 환율 예측에 미치는 영향을 실증적으로 분석하였다.

모든 수치형 변수에 대해서는 스케일 차이를 보정하고 학습 안정성을 높이기 위해 Min-Max 스케일링을 적용하였으며, 최소값은 0, 최대값은 1로 정규화하였다.

예측 모델은 시계열 딥러닝 모델(LSTM, GRU, CNN-LSTM, CNN-GRU)에 감정 점수 변수를 결합하여 총 14가지 조합으로 실험을 진행하였다. 감정 점수는 뉴스와 댓글로부터 각각 산출되었으며, 구성된 변수 조합은 다음 표와 같다:

|  |  |
| --- | --- |
| 분류 | 설명 |
| 뉴스 기반 감정 분석 | 긍정/부정/중립 점수 (3분류) |
| 긍/부/중 점수 + 감정 score ≥ 0.95 |
| 긍/부/중 점수 + 감정 score ≥ 0.90 |
| 긍정/부정 점수 (2분류) |
| 긍/부 점수 + 감정 score ≥ 0.95 |
| 긍/부 점수 + 감정 score ≥ 0.90 |
| 댓글 기반 감정 분석 | 긍정/부정 점수 (탐욕/공포 기준, 2분류) |
| 긍/부 점수 + 댓글 개수 변수 포함 |
| 감정 score ≥ 0.90 점수만 포함 |
| 감정 score ≥ 0.90 + 댓글 개수 변수 포함 |
| 혼합 | **뉴스 감성 분석(긍부정중립) + 댓글 감성 분석(긍부정)** |
| **뉴스 감성 분석(긍부정) + 댓글 감성 분석(긍부정)** |
| **뉴스 감성 분석(긍부정중립) + 댓글 감성 분석(긍부정) + 댓글 갯수** |
| **뉴스 감성 분석(긍부정) + 댓글 감성 분석(긍부정) + 댓글 갯수** |

이러한 조합을 통해 감정 점수의 존재 여부, 분류 방식(2분류 vs 3분류), 신뢰도(score threshold), 그리고 보조 지표(댓글 수)의 포함 여부가 환율 예측에 미치는 영향을 다각도로 분석하였다.

* 1. *Prediction Performance*

본 연구에서는 환율 예측을 위해 LSTM, GRU, CNN-LSTM, CNN-GRU의 네 가지 딥러닝 모델을 활용하였다. 전체 데이터 중 80%를 학습용(training set)으로 사용하여 모델을 학습하고, 하이퍼파라미터 튜닝을 통해 최적의 모델 구조를 도출하였다. 나머지 20%는 테스트용(test set)으로 활용하여 미래 시점에 대한 예측 성능을 평가하고, 모델의 일반화 능력을 검증하였다.

모델 학습의 일관성과 재현 가능성을 확보하기 위해, 모든 실험에서 난수 고정(seed) 값을 42로 설정하였다. 이를 통해 반복 실험 시 동일한 조건에서 결과를 비교할 수 있도록 하였다.

T시점의 환율을 예측하기 위해 과거 5일의 데이터를 바탕으로 다음 날의 환율을 예측하는 모델링을 수행하였다. (window\_size = 5)

표 3. 공통 하이퍼파라미터

|  |  |
| --- | --- |
| 하이퍼파라미터 | 값 |
| Window Size | 5 |
| Epochs | 50 |
| Bath Size | 16 |
| Optimizer | Adam |
| Loss Function | Mean Squared Error |
| Early Stopping | monitor='loss' |
| patience=5 |
| restore\_best\_weights=True |
| Output Activation | Linear |
| Target Scaling | MinMaxScaler |
| Seed | 42 |

|  |  |  |
| --- | --- | --- |
| **Algorithm** | **Parameter** | **Value** |
| **LSTM** | Unit Numbers | 50 |
| Dense Units | 1 |
| Dropout Ratio | 0.2 |
| **GRU** | Unit Numbers | 50 |
| Dense | 1 |
| Dropout Ratio | 0.2 |
| **CNN-LSTM** | Unit Numbers | 50 |
| Dense Units | 2 |
| Dropout Ratio | 0.2 |
| filters | 256 |
| kernel\_size | 3 |
| activation | ReLU |
| Pooling | MaxPooling1D(pool\_size=2) |
| Dense Units | 1 |
| **CNN-GRU** | Unit Numbers | 50 |
| Dense Units | 2 |
| Dropout Ratio | 0.2 |
| filters | 256 |
| kernel\_size | 3 |
| activation | ReLU |
| Pooling | MaxPooling1D(pool\_size=2) |
| Dense Units | 1 |

감정 분석 변수를 포함한 모델이 그렇지 않은 모델보다 전반적으로 더 낮은 예측 오차를 보이며 우수한 성능을 나타냈다. 특히, GRU 모델이 대부분의 조합에서 가장 낮은 MSE 및 MAPE를 기록하며 예측 정확도 측면에서 가장 안정적인 성능을 보였다. 그 중 Socre 0.9이상 댓글 감정 분석을 조합한 GRU 모델이 mape 0.557374로 가장 성능이 좋게 나왔다.

감정 분석 변수 중에서는 뉴스 데이터를 기반으로 한 긍·부정 점수(2분류) 또는 댓글 데이터를 기반으로 한 감정 점수 중 신뢰도(score) 0.9 이상인 경우가 예측 성능을 가장 크게 향상시키는 것으로 나타났다. 예를 들어, 감정 분석을 전혀 사용하지 않은 GRU 모델의 MSE는 391.89, MAPE는 1.41로 나타났으나, 뉴스 감성 점수(2분류)를 포함한 경우 MSE는 65.01, MAPE는 0.57까지 감소하였다. 또한, 댓글 감정 점수(score > 0.9)를 포함한 GRU 모델의 MSE는 70.98, MAPE는 0.56으로 매우 우수한 예측력을 보였다.

반면, CNN-LSTM과 CNN-GRU 기반 모델은 상대적으로 예측 오차가 큰 편이었으며, 특히 뉴스·댓글 감정 변수를 포함하더라도 예측 정확도 개선 효과가 제한적으로 나타났다. 이는 CNN 구조가 시계열 데이터의 장기적인 흐름을 포착하는 데 상대적으로 비효율적일 수 있음을 시사한다.

또한, 뉴스와 댓글 감정 점수를 동시에 포함한 조합도 실험에 포함하였으며, 일부 경우에서는 단일 감정 정보만을 사용했을 때보다 더 나은 성능을 보였다. 그러나 모든 조합에서 일관된 개선 효과를 보이지는 않았으며, 단일한 고신뢰 감정 점수(score ≥ 0.9)를 사용하는 것이 더 안정적인 성능을 제공하는 것으로 분석되었다.

|  |  |  |  |
| --- | --- | --- | --- |
| Combination | Model | MSE | MAPE |
| w/o sentiment score | LSTM | 154.582199 | 0.862627 |
| GRU | 391.894196 | 1.414386 |
| CNN-LSTM | 1231.639771 | 2.51515 |
| CNN-GRU | 732.919067 | 1.890148 |
| 뉴스 감정(3분류) | LSTM | 490.783325 | 1.627882 |
| GRU | 71.056755 | 0.57595 |
| CNN-LSTM | 1203.92981 | 2.541264 |
| CNN-GRU | 501.083466 | 1.615509 |
| 뉴스 감정(3분류, score>0.95) | LSTM | 383.701355 | 1.412983 |
| GRU | 99.044571 | 0.679252 |
| CNN-LSTM | 464.975739 | 1.53446 |
| CNN-GRU | 472.956818 | 1.563344 |
| 뉴스 감정(3분류, score>0.90) | LSTM | 614.722107 | 1.831132 |
| GRU | 54.050175 | 0.481272 |
| CNN-LSTM | 1521.192383 | 2.866176 |
| CNN-GRU | 1051.240479 | 2.357519 |
| 뉴스 감정(2분류) | LSTM | 211.396454 | 1.049603 |
| GRU | 975.06543 | 2.473251 |
| CNN-LSTM | 800.985352 | 2.065083 |
| CNN-GRU | 820.174866 | 2.077493 |
| 뉴스 감정(2분류, scroe > 0.95) | LSTM | 218.518158 | 1.037139 |
| GRU | 98.71032 | 0.677744 |
| CNN-LSTM | 678.608887 | 1.83707 |
| CNN-GRU | 476.040558 | 1.524749 |
| 뉴스 감정(2분류, scroe > 0.90) | LSTM | 227.091049 | 1.103213 |
| GRU | 132.248749 | 0.743927 |
| CNN-LSTM | 638.762207 | 1.795203 |
| CNN-GRU | 1177.116943 | 2.579644 |
| 댓글 감성 (2분류) | LSTM | 1154.913086 | 2.413401 |
| GRU | 89.624794 | 0.606184 |
| CNN-LSTM | 1062.743774 | 2.402438 |
| CNN-GRU | 1109.994507 | 2.449066 |
| 댓글 감성(2분류) + 댓글 갯수 | LSTM | 235.110687 | 1.100125 |
| GRU | 102.392662 | 0.670695 |
| CNN-LSTM | 1330.480347 | 2.678723 |
| CNN-GRU | 515.819885 | 1.611753 |
| 댓글 감성 (2분류, score > 0.9) | LSTM | 194.402954 | 0.964966 |
| GRU | 82.081612 | 0.586493 |
| CNN-LSTM | 1336.286011 | 2.612379 |
| CNN-GRU | 603.572205 | 1.744603 |
| 댓글 감성 (2분류, score > 0.9) + 댓글 갯수 | LSTM | 339.607697 | 1.332964 |
| GRU | 77.145309 | 0.636674 |
| CNN-LSTM | 922.363647 | 2.165729 |
| CNN-GRU | 870.243652 | 2.046372 |
| 뉴스 감성 분석(긍부정중립) + 댓글 감성 분석(긍부정) | LSTM | 476.359436 | 1.551748 |
| GRU | 101.018295 | 0.715507 |
| CNN-LSTM | 1122.296265 | 2.41336 |
| CNN-GRU | 491.57312 | 1.632307 |
| 뉴스 감성 분석(긍부정) + 댓글 감성 분석(긍부정) | LSTM | 156.567551 | 0.87081 |
| GRU | 222.289429 | 1.174175 |
| CNN-LSTM | 757.162964 | 1.968055 |
| CNN-GRU | 1073.173218 | 2.32297 |
| 뉴스 감성 분석(긍부정중립) + 댓글 감성 분석(긍부정) + 댓글 갯수 | LSTM | 304.60379 | 1.246441 |
| GRU | 204.692917 | 0.870027 |
| CNN-LSTM | 775.353394 | 1.99209 |
| CNN-GRU | 403.933228 | 1.42059 |
| 뉴스 감성 분석(긍부정) + 댓글 감성 분석(긍부정) + 댓글 갯수 | LSTM | 925.945679 | 2.225798 |
| GRU | 70.657043 | 0.559794 |
| CNN-LSTM | 1016.737122 | 2.350776 |
| CNN-GRU | 972.670227 | 2.275009 |

이러한 결과는 시장 심리를 반영한 감정 점수가 환율 예측 성능 향상에 기여할 수 있음을 실증적으로 보여주며, 특히 GRU와 같은 구조가 감정 정보와 시계열 데이터를 효과적으로 통합할 수 있음을 시사한다.

GRU

|  |  |  |  |
| --- | --- | --- | --- |
| Combination | Model | MSE | MAPE |
| w/o sentiment score | GRU | 391.894196 | 1.414386 |
| 뉴스 감정(3분류) | GRU | 71.056755 | 0.57595 |
| 뉴스 감정(3분류, score>0.95) | GRU | 99.044571 | 0.679252 |
| 뉴스 감정(3분류, score>0.90) | GRU | 54.050175 | 0.481272 |
| 뉴스 감정(2분류) | GRU | 975.06543 | 2.473251 |
| 뉴스 감정(2분류, scroe > 0.95) | GRU | 98.71032 | 0.677744 |
| 뉴스 감정(2분류, scroe > 0.90) | GRU | 132.248749 | 0.743927 |
| 댓글 감성 (2분류) | GRU | 89.624794 | 0.606184 |
| 댓글 감성(2분류) + 댓글 갯수 | GRU | 102.392662 | 0.670695 |
| 댓글 감성 (2분류, score > 0.9) | GRU | 82.081612 | 0.586493 |
| 댓글 감성 (2분류, score > 0.9) + 댓글 갯수 | GRU | 77.145309 | 0.636674 |
| 뉴스 감성 분석(긍부정중립) + 댓글 감성 분석(긍부정) | GRU | 101.018295 | 0.715507 |
| 뉴스 감성 분석(긍부정) + 댓글 감성 분석(긍부정) | GRU | 222.289429 | 1.174175 |
| 뉴스 감성 분석(긍부정중립) + 댓글 감성 분석(긍부정) + 댓글 갯수 | GRU | 204.692917 | 0.870027 |
| 뉴스 감성 분석(긍부정) + 댓글 감성 분석(긍부정) + 댓글 갯수 | GRU | 70.657043 | 0.559794 |

# Conclusions

본 연구는 뉴스 기사 및 댓글 데이터를 기반으로 감정 분석을 수행하고, 이를 딥러닝 기반 시계열 예측 모델에 통합하여 USD/KRW 환율을 예측하는 방법을 제안하였다. 특히 감정 분석 결과를 정량화하여 입력 변수로 활용하고, LSTM, GRU, CNN-LSTM, CNN-GRU의 네 가지 딥러닝 모델을 비교함으로써 감정 분석의 효과성과 모델별 예측 성능을 실증적으로 분석하였다.

연구 결과, 감정 분석 변수를 포함한 모델이 전통적인 경제 지표만을 활용한 모델보다 전반적으로 더 우수한 예측 성능을 보였다. 특히 GRU 모델은 대부분의 실험 조합에서 가장 낮은 MSE 및 MAPE를 기록하며, 시계열 데이터와 감정 점수의 통합 예측에 가장 효과적인 구조로 나타났다.

감정 분석 방식 측면에서는, 뉴스 데이터를 기반으로 한 3분류(긍정/부정/중립) 고신뢰(score ≥ 0.9) 감정 점수 조합이 예측 정확도 향상에 가장 크게 기여하였다. 반면, CNN 기반 모델(CNN-LSTM, CNN-GRU)은 시계열 예측에서 상대적으로 낮은 성능을 보여, 환율과 같은 시계열적 흐름이 중요한 변수에서는 RNN 계열 모델의 적합성이 더 높은 것으로 판단된다.

이러한 결과는 기존 환율 예측 연구가 주로 거시경제 지표에 의존해 왔던 한계를 넘어, 시장 참여자의 심리를 반영한 비정형 텍스트 데이터의 활용 가능성을 제시한다. 특히 본 연구는 뉴스와 댓글의 감정 정보를 정량화하여 모델에 효과적으로 통합함으로써, 시장 반응과 기대를 보다 정교하게 반영하는 예측이 가능함을 확인하였다.

향후 연구에서는 더욱 다양한 비정형 데이터 소스를 확대하고, Transformer 기반의 시계열 모델 등 최신 구조를 적용함으로써 감정 기반 예측 모델의 정확도와 실용성을 한층 더 향상시킬 수 있을 것으로 기대된다.

# Data Availability

The data can be provided on request.

# Declaration of Competing Interest

The authors declare no conflict of interest.

# CRediT authorship contribution statement

**Jaeheung Park:** Conceptualization, Software, Validation, Data Curation, Writing Original Draft, Preparation.

**Kyungwon Kim:** Methodology, Preparation, Software, Writing Review and Editing, Visualization, Supervision.

# References

Adadi, A., and M. Berrada. 2018. 'Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)', *IEEE Access, Access, IEEE*, 6: 52138-38.

Alsubari, Saleh Nagi, Sachin N. Deshmukh, Mosleh Hmoud Al-Adhaileh, Fawaz Waselalla Alsaade, and Theyazn H. H. Aldhyani. 2021. 'Development of Integrated Neural Network Model for Identification of Fake Reviews in E-Commerce Using Multidomain Datasets', *Applied Bionics and Biomechanics*, 2021: 5522574.

Cha, Gi-Wook, Hyeun-Jun Moon, and Young-Chan Kim. 2021. 'Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables', *International journal of environmental research and public health*, 18: 8530.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. San Francisco, California, USA: Association for Computing Machinery.

Fleck, Anna. 2023. 'Fare Play: Who's Riding Public Transit?', Statista, Accessed 01.07. <https://www.statista.com/chart/30882/share-of-commuters-that-take-public-transport/>.

Michie, D., D. Spiegelhalter, and Charles Taylor. 1999. 'Machine Learning, Neural and Statistical Classification', *Technometrics*, 37.

Molnar, Christoph. 2023. *Interpretable machine learning. A Guide for Making Black Box Models Explainable* (Lulu. com).

Office, Seoul City Transportation. 2023. "Seoul Transportation in 2022." In, edited by Futuristic Advanced Transportation Division.

Shustanov, Alexander, and Pavel Yakimov. 2017. 'CNN Design for Real-Time Traffic Sign Recognition', *Procedia Engineering*, 201: 718-25.

Viswavandya, Meera, Shashwat Patel, and Kaushik Sahoo. 2021. 'ANALYSIS AND COMPARISON OF MACHINE LEARNING APPROACHES FOR TRANSMISSION LINE FAULT PREDICTION IN POWER SYSTEMS', *Journal of Research in Engineering and Applied Sciences*, 6: 24-31.

권효승, 최창준, 정효석, 송재인, 강민희, and 황기연. 2020. 'GRU 를 이용한 서울시 지하철 구간별 혼잡도 예측 모델 연구', *한국 ITS 학회 학술대회*: 185-90.

김규진, 박성환, 최성훈, and 김양중. 2024. '딥러닝 분석을 이용한 지하철역 혼잡도 변화의 동적 분석 및 예측 모델에 관한 연구', *한국통신학회 학술대회논문집*, 2024: 1192-93.

김승준. 2016. "서울시 지하철의 혼잡비용 산정과 정책적 활용방안." In, 1-22.

김재익. 2013. '아침 첨두시간대 지하철 이용수요의 결정요인에 관한 연구 - 대구 지하철 역세권 토지이용을 중심으로', *교통연구*, 20: 15-25.

김진수. 2016. '빅데이터 분석을 이용한 지하철 혼잡도 예측 및 추천시스템', *디지털융복합연구*, 14: 289-95.

성현곤. 2017. '서울시 대중교통 수단별 월별 이용수요의 변동에 영향을 미치는 요인 분석', *국토계획*, 52: 81-96.

이경재, 김수재, 문형택, 한재윤, and 추상호. 2020. '심층신경망 모형을 활용한 대중교통 이용자의 환승시간 추정에 관한 연구', *한국ITS학회논문지*, 19: 32-43.

이상준, 신성일, 이성희, and 윤성진. 2021. '승객 승하차 이동 및 대기행태를 고려한 도시철도 승강장 점유면적 추정방안', *대한교통학회지*, 39: 721-35.

이상준, 신성일. 2020. '교통카드 자료를 활용한 지하철 혼잡도 개선 연구 : Early Bird 정책대안을 중심으로', *한국IT서비스학회지*, 19: 125-38.

이선하, 천춘근, 정병두, 유병영, and 김은지. 2015. '철도 이용객 정보제공 효과평가 방법론 연구 -승강장의 혼잡상황을 고려한 Gate Metering 사례 연구 중심으로', *한국ITS학회 논문지*, 14: 50-62.

이정훈, and 정헌영. 2018. '분위 회귀를 활용한 기상조건이 대중교통 수단별 통행량에미치는 영향에 대한 연구', *국토계획*, 53: 95-106.

이호, and 최진경. 2015. '대중교통카드 자료를 활용한 도시철도 승강장 혼잡도 추정 알고리즘 개발', *한국철도학회논문집*, 18: 270-77.

장진영, 김채원, and 박민서. 2023. '머신러닝 기반 2호선 출퇴근 시간대 지하철 역사 내 혼잡도 예측', *문화기술의 융합*, 9: 145-50.

천예은, 김세빈, 이자윤, and 우지환. 2021. '설명 가능한 AI 기술을 활용한 신용평가 모형에 대한 연구', *한국데이터정보과학회지*, 32: 283-95.