글로벌 이벤트와 감성 융합을 활용한 한국의 딥러닝 단기 환율 예측

Global Events and Sentiment Fusion for Deep Learning–Based Short-Term Exchange Rate Prediction in Korea

Jaeheung Parka, Kyungwon Kima,*∗*

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

# 요약

# *본 연구는 원/달러 환율의 단기 예측 성능을 향상시키기 위해 거시·금융시장 지표와 더불어 뉴스 기반 이벤트(GDELT) 및 감성 요약 지표를 통합한 딥러닝 모델을 제안하였다. LSTM, GRU, CNN-LSTM, CNN-GRU 네 가지 아키텍처를 대상으로 다양한 입력 조합과 시차 구간(lookback=5~90일)을 실험한 결과, 이벤트·감성 통합 입력과 CNN-LSTM(lookback=20)의 조합이 최적 성능을 달성하였다(MSE=47.71, MAPE=0.39%). 변수 중요도 분석(Permutation Importance)을 통해 금·알루미늄·구리 등 원자재 가격, 미국 정책금리와 국채금리, KOSPI/KOSDAQ 지수 및 거래량 등이 환율 변동의 핵심 설명력을 가지는 것으로 확인되었으며, 뉴스 요약의 감성 변동성 역시 의미 있는 신호로 작용하였다. 시각화 결과, CNN-LSTM 기반 예측치는 실제 환율과 유사한 흐름을 포착하며 단기 방향성을 효과적으로 설명하였다. 본 연구는 사건·감성 정보의 활용이 환율 예측의 정확도를 높일 수 있음을 실증적으로 보여주며, 향후 금융시장 예측에 있어 텍스트 기반 신호의 중요성을 뒷받침한다.*

# *중심어 : 환율, 시계열 예측, 딥러닝, 글로벌 이벤트, 뉴스 감성 분석*

# Abstract

*This study proposes a deep learning framework for enhancing short-term forecasting of the KRW/USD exchange rate by integrating macro-financial indicators with news-based event (GDELT) and sentiment summary variables. Four architectures—LSTM, GRU, CNN-LSTM, and CNN-GRU—were evaluated across multiple input settings and lookback windows (5–90 days). The combination of event–sentiment integration and CNN-LSTM with a 20-day lookback achieved the best performance (MSE=47.71, MAPE=0.39%). Permutation importance analysis revealed that commodity prices (gold, aluminum, copper), U.S. policy and bond yields, and domestic equity indices (KOSPI/KOSDAQ) exerted the strongest explanatory power on exchange rate fluctuations, while sentiment volatility in news summaries also provided significant predictive signals. Visualization of the best-performing model demonstrated that CNN-LSTM predictions closely tracked actual exchange rate dynamics, effectively capturing short-term market direction. These findings provide empirical evidence that incorporating event and sentiment information can improve forecasting accuracy, underscoring the role of text-based signals in financial market prediction.*

*Keywords:* Exchange Rate, Time Series Forecasting, Deep Learning, Global Events, News Sentiment

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

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

# Ⅰ. 서론

환율(Exchange rate)은 외국 통화 한 단위를 획득하기 위해 자국 통화를 얼마나 지불해야 하는지를 나타내는, 즉 자국 통화와 외국 통화 간의 교환 비율을 의미한다[1]. 이는 두 나라 통화의 상대적 가치를 보여주는 지표로, 오늘날 거의 모든 경제 활동이 국제 거래와 밀접하게 연계되어 있기 때문에 개인의 일상생활 뿐 아니라 기업 경영, 국가 경제 운영에 이르기까지 필수적인 정보로 기능한다[2]. 특히 원/달러 환율은 한국 대외경제에서 핵심적인 역할을 한다. 환율 변동은 수출입 가격, 외국인 투자, 물가 수준, 중앙은행의 통화정책 등 다양한 거시경제 변수와 밀접하게 연관되어 있으며, 환율의 향후 변화를 예측하는 것은 국가 경제의 안정성과 기업의 재무 전략 수립에 있어 중요한 과제로 인식된다. 그러나 환율은 경제 펀더멘털뿐만 아니라 지정학적 리스크, 글로벌 자본 흐름, 시장 심리 등 복합적인 요인의 영향을 받는다. 이로 인해 환율 변동은 종종 ‘랜덤워크(Random Walk)’로 설명되며, 이는 술에 취한 사람이 어디로 걸어갈지 예측하기 어렵듯 환율의 다음 시점을 정확히 전망하기 힘들다는 비유적 표현이다[3].

환율은 외환시장의 수요와 공급에 의해 결정되며, 장기적으로는 물가 수준이나 생산성 변화와 같은 경제 여건이 통화 가치에 영향을 미친다. 중기적으로는 대외거래와 거시경제정책이 주요 요인으로 작용하고, 단기적으로는 시장 참가자들의 기대, 주변국 환율 변동, 각종 뉴스에 의해 크게 좌우된다. 특히 시장 참가자들의 기대가 환율 상승 혹은 하락 쪽으로 쏠리면 자기실현적(Self-fulfilling) 거래를 통해 실제 환율 변동을 초래하기도 한다. 예컨대 다수가 환율 상승을 예상할 경우, 환율이 오르기 전에 외환을 선매입하려는 수요가 증가하여 실제로 환율 상승을 유발하게 된다. 또한 뉴스와 같은 정보 역시 시장 참가자들의 기대를 변화시켜 단기 환율 변동에 중요한 영향을 미친다. 실제로 2010년 5월 천안함 침몰조사 결과 발표로 지정학적 리스크가 부각되자 원/달러 환율이 단기간 급등한 사례가 있다[4].

이와 같은 환경에서 환율의 경로를 정밀하게 예측하는 것은 정책당국의 통화·재정정책 수립, 기업의 수출입 가격전략 조정, 금융기관의 리스크 관리에 필수적이다. 기존 국내 연구들은 이러한 필요성을 강조해왔다.

김우석·한규식(2021)은 팬데믹 이후 환율의 구조적 변화를 분석하며 변동성 확대에 따른 예측 필요성을 제시하였고[14], 오인정·김우주(2022)는 코로나19 구간을 대상으로 SARIMA와 ARDL 모형을 비교하여 국면별로 원/달러 예측모형의 성능이 달라진다는 점을 보여주었다[15].

그러나 이러한 연구들은 대부분 종가 기반의 시계열 모형 비교 또는 구조변화 탐색에 초점을 맞추었으며, 이벤트 데이터와 감성 분석 지표를 결합하여 환율을 예측한 사례는 드물다. 이에 본 연구는 기존 선행연구의 거시경제적·구조적 요인을 포괄하면서, 나아가 뉴스 감성 점수와 GDELT 기반 이벤트 지표를 통합한 융합적 접근을 시도한다. 이를 통해 환율 변동성의 구조적 요인과 심리적 요인을 동시에 반영하는 정교한 예측 모델을 제시하고자 한다.

이러한 맥락에서 시장 참여자들의 심리와 뉴스에 담긴 감정을 정량화해 반영하는 감성 분석(Sentiment Analysis)은 환율 예측의 새로운 단서를 제공할 수 있다. 따라서 본 연구는 감성 지표를 분석하여 이를 딥러닝 기반 시계열 예측 모델에 결합함으로써, 원/달러 환율 예측의 정확도를 향상시킬 수 있는 가능성을 탐색하고자 한다.

# Ⅱ. 문헌연구

환율 예측 연구들은 전통적으로 거시경제 변수와 금융시장 변수를 주요 독립변수로 설정하였다.  
Plakandaras et al.(2015)은 1999년부터 2011년까지의 데이터를 바탕으로 원자재(19개), 금속(10개), 주가지수(7개), 무역지수(3개), 금리(11개), 거시경제 지표(11개), 환율(8개) 등 광범위한 변수를 활용하였다. 다양한 모형을 비교한 결과, EEMD 기반 비선형 모형이 일간 및 월간 환율 예측에서 우수한 성능을 보였다[5].  
국내 연구로는 임현욱 외(2021)가 연합인포맥스와 Bloomberg 데이터를 사용하여 원/달러 환율을 예측하였다. 이들은 1년·10년 만기의 원화 IRS 금리, KTB 수익률, 본드 스왑 스프레드 등 금리 및 채권시장 변수를 중심으로 구성하였으며, ANN 모델이 평균 Hit Ratio 50.96%로 가장 우수한 성과를 보였다[6]. Cao et al.(2020)은 유가와 금값 같은 실물자산 변수 뿐 아니라 통화공급(M1, M2), CPI, PPI, 산업생산, 기준금리, 인플레이션율, 무역수지, 정책 불확실성 지수 등 거시경제 지표를 결합하였다. 이들은 제안한 DC-LSTM 모형이 CNN, LSTM, SVR, ARIMA보다 낮은 예측 오차(MAE 0.0145)를 기록했다고 보고하였다[7].  
또한 Qureshi(2025)는 USD/INR 환율을 대상으로 금리, 인플레이션율, GDP 성장률, 외환보유액 등 거시경제 펀더멘털 변수를 활용하였다. 다양한 머신러닝 모형을 비교한 결과, Random Forest가 RMSE 0.029로 가장 우수한 성과를 나타냈다[8]. 마지막으로 Wang et al.(2021)은 USD/CNY 환율 뿐만 아니라 나스닥, 다우존스, 상하이 종합지수, 항셍지수 등 글로벌 주가지수를 독립변수로 결합하였다. CNN, RNN, LSTM 등 여러 모델을 비교한 결과, CNN-TLSTM 모형이 MAPE 0.18945로 가장 높은 예측력을 보였다[9]. 종합하면, 기존 환율 예측 연구들은 공통적으로 금리, 물가, 무역수지, 주가지수, 원자재 가격 등 거시경제 및 금융 변수를 중심으로 환율 변동을 설명하고자 했음을 확인할 수 있다.

기존 선행연구 대비 본 연구의 가장 큰 차별성은 전통적 변수군에 외생적 사건을 반영하는 이벤트 기반 지표를 통합한 점에 있다. 이를 위해 GDELT(Global Database of Events, Language, and Tone) 데이터베이스를 활용하였으며, 이는 전 세계 언론(인쇄·방송·웹 기반)을 실시간으로 모니터링하여 사건 유형, 위치, 감정 톤, 인물 및 조직, 문서 수 등 다양한 속성으로 사건을 정량화하는 데이터베이스이다[16]. GDELT의 데이터는 전 세계 사회적 사건을 정량적 시계열 데이터로 전환할 수 있다는 점에서 기존 거시·금융 예측 변수의 한계를 보완할 수 있는 강력한 도구로 주목받고 있다.

경제 및 금융 분야에서도 GDELT 기반 이벤트 데이터를 예측 변수로 활용한 사례가 있다. 예컨대 ARIMA 모형에 GDELT 데이터를 추가하여 환율 예측 정확도를 약 1% 향상시키고, Gradient Boosting 기반 분류 모델에서는 정확도를 6% 상승시킨 연구[17]와, 유럽 채권시장 분석에서 GDELT 지표가 nowcasting 모델의 성능을 향상시킨 사례가 있다[18]. 또한 미국 주식시장 예측에서, GDELT 지표는 전통적 거시 경제 모형을 능가하는 설명력 및 예측 성능을 제공했다는 실증 결과도 보고되었다[19].

본 연구는 뉴스 기사와 온라인 댓글이라는 두 가지 비정형 텍스트 데이터를 대상으로 감성 분석을 수행하였다. 감성 분석은 자연어 처리(NLP)의 하위 분야로, 텍스트에 내재된 정서적 톤을 식별하고 분류하는 작업을 의미한다[20].

기존 연구에서는 규칙 기반 접근이나 전통적 머신러닝 기법이 주로 활용되었으나, 도메인 의존성, 언어 확장성, 은유·풍자와 같은 복잡한 뉘앙스 처리의 한계가 지속적으로 지적되었다[21]. 이러한 제약을 극복하기 위해 최근에는 Transformer 기반 딥러닝 모델이 적극 도입되었다. 특히 BERT 계열 모델은 양방향 문맥 정보를 활용하여 단어 의미를 정교하게 파악하고, 적은 데이터로도 파인튜닝이 가능하다는 장점이 있다.

금융 분야에서는 FinBERT와 같이 금융 특화 언어모델이 등장하여 기존 머신러닝 접근 대비 성능이 크게 향상되었으며[22], FiQA·Financial PhraseBank와 같은 벤치마크에서도 Transformer 기반 모델들이 전통적 방법을 압도하는 성능을 기록하였다[23]. 더 나아가 최근에는 GPT-4o와 같은 대규모 언어모델(LLM)이 프롬프트 튜닝이나 few-shot 학습만으로도 FinBERT와 동등하거나 그 이상의 성능을 보이는 것으로 보고되고 있다[24][25]. 이는 금융 뉴스·리포트 분석과 같은 복잡한 텍스트 처리에서 LLM이 높은 적응력과 확장성을 제공함을 보여준다.

이러한 흐름을 반영하여, 본 연구는 전통적인 규칙 기반·머신러닝 접근 대신 GPT-4o-mini 기반 LLM 감성 분석 절차를 설계하였다.

최근 금융시장 예측에서는 전통적인 거시·금융 변수 뿐 아니라 시장 참여자의 심리 변수를 반영하기 위해 감성 분석이 적극적으로 활용되고 있다.  
Mohan et al.(2019)은 Apple, Amazon, Microsoft 등 S&P500 주요 기업의 종가 정보와 함께 국제 뉴스 기사에서 추출한 긍·부정 감정 점수를 독립변수로 사용하였다. 분석 결과, 감정 변수를 포함한 RNN 모델이 MAPE 2.03으로 가장 우수한 성능을 보였다[10]. Jing et al.(2021)은 Shanghai Stock Exchange 데이터와 Eastmoney.com 투자자 게시글의 감정 점수를 CNN으로 분류하고, 이를 주가 기술적 지표와 함께 LSTM 모델에 입력하였다. 그 결과 CNN-LSTM 모델이 평균 MAPE 0.0449로 가장 뛰어난 예측력을 보여, 투자자 심리 반영의 효과성을 입증하였다[11]. 또한 정가연 외(2024)는 네이버 증권 뉴스, 한국은행 API, Naver DATALAB 데이터를 활용하여 코스피, S&P500, 유가, 금값, 원/엔 환율, 금리 등 거시·금융 변수와 함께 KLUE-BERT 기반 뉴스 감정 점수, 코스피 검색 빈도수 등 시장 심리 변수를 결합하였다. 실험 결과, CNN-GRU 모델이 MAPE 1.38%로 가장 낮아 감성 분석의 유용성을 보여주었다[12]. 이러한 연구들은 전통적인 금융 변수에 뉴스 및 게시글 감정 점수를 결합하는 융합적 독립변수 구성이 주가 예측 정확도를 높일 수 있음을 시사한다.

환율 예측에도 최근에는 감성 분석을 접목한 연구가 시도되고 있다. Ding et al.(2024)은 Investing.com과 ForexEmpire의 뉴스 및 댓글 데이터를 기반으로 감정 점수를 산출하였다. 이들은 해당 감성 지표를 USD/CAD, USD/MXN, USD/CNY, USD/JPY, USD/KRW 등 주요 교차환율, 원자재 가격(유가, 금, 구리, 옥수수 등), 글로벌 지수(S&P500, EuroStoxx50, VIX), 채권수익률(미국·유로존 10년물)과 같은 거시·금융 변수와 결합하였다. 다양한 모형을 비교한 결과, PSO-LSTM 모델이 RMSE 0.0958로 가장 높은 성능을 기록하였다[13]. 이러한 결과는 환율 예측에서도 단순한 거시경제 지표만으로 설명하기 어려운 단기 변동성을 뉴스 및 댓글에 담긴 시장 심리 변수가 효과적으로 보완할 수 있음을 보여준다.

<표 1>에 정리된 선행연구를 종합하면, 다음과 같은 시사점을 도출할 수 있다. 첫째, 환율 예측에서는 여전히 금리, 물가, 무역수지, 원자재 가격, 주가지수 등 거시경제 및 금융 변수가 핵심적 설명 변수로 기능하고 있다. 둘째, 최근 연구들은 뉴스, 투자자 게시글, 댓글 등 비정형 텍스트 데이터에서 추출한 감정 점수를 독립변수로 추가하여 예측 성능을 향상시키고 있으며, 이는 금융시장 예측에서 시장 심리 변수의 중요성을 뒷받침한다. 셋째, 감성 분석을 결합한 환율 예측 연구는 아직 초기 단계에 있으나, 뉴스·댓글 감성과 같은 심리 변수의 확장적 활용 가능성이 크다는 점에서 향후 연구 방향성을 제시한다. 따라서 본 연구는 이러한 시사점을 바탕으로, 거시경제 변수와 함께 GDELT 이벤트 톤, 뉴스 감성 분석을 통합한 딥러닝 기반 원/달러 환율 예측 모델을 제안한다. 이를 통해 환율 변동성을 설명할 수 있는 새로운 변수를 발굴하고, 기존 연구 대비 예측 정확도를 향상시키고자 한다.

**<표 1> 환율 예측 및 주가 예측과 감성 분석을 활용한 선행연구 분석 방향과 성능 요약**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 연구 목적 | 연구자(연도) | 데이터 소스 | 데이터 기간 | 주요 독립변수 | 방법론 | 성능 |
| **환율 예측** | Plakandaras et al.(2015)  [5] | 미기재 | 1999~2011 | 원자재(19), 금속(10), 주가지수(7), 금리(11), 무역·거시지표 등 | ARIMA, GARCH, AR-NN, MARS-NN, SVR, EEMD 결합 | 일간: EEMD-MARS-SVR, MARS-NN 우수 / 월간: EEMD-AR-SVR 우수 |
| 임현욱 외(2021)  [6] | 연합인포맥스, Bloomberg | 2008~2020 | IRS 금리, KTB 수익률, 본드 스왑 스프레드 | ANN, LR, DT | ANN 평균 Hit Ratio 50.96% |
| Cao et al.(2020)  [7] | 미기재 | 2016.6~2019.4 | 유가, 금값, CPI, PPI, 산업생산, 기준금리, 무역수지, 정책 불확실성 지수 등 | ARIMA, SVR, CNN, LSTM, DC-LSTM | DC-LSTM MAE 0.0145 |
| Qureshi(2025)  [8] | FRED, IMF, World Bank 등 | 2015~2023 | 금리, 인플레이션율, GDP 성장률, 외환보유액 | ARIMA, RW, RF, SVM | RF RMSE 0.029 |
| Wang et al.(2021)  [9] | Wind DB | 2006~2020 | USD/CNY, 나스닥, 다우, 상하이, 항셍 등 지수 | MLP, CNN, RNN, LSTM, CNN-LSTM, CNN-TLSTM | CNN-TLSTM MAPE 0.18945 |
| **주가 예측 + 감성분석** | Mohan et al.(2019)  [10] | 해외 뉴스 웹사이트 | 2013~2017 | S&P500 기업 종가, 뉴스 감정 | ARIMA, Prophet, RNN | RNN-pp MAPE 2.03 |
| Jing et al.(2021)  [11] | SSE, Eastmoney.com | 2017~2019 | 기술적 지표 , 게시글 감정 | CNN+LSTM | MAPE 0.0449 |
| 정가연 외(2024)  [12] | 네이버뉴스, KOSPI, DATALAB, 한은 API | 2021~2023 | 코스피·S&P500, 유가, 금값, 원/엔 환율, 금리, 검색 빈도, 감정 점수 | LSTM, GRU, CNN-LSTM, CNN-GRU | CNN-GRU + KLUE-BERT 감정 MAPE 1.38% |
| **환율 예측 + 감성분석** | Ding et al.(2025)  [13] | Investing.com, ForexEmpire | 2016~2024 | 뉴스·댓글 감정, 교차환율, 원자재, 글로벌 지수, 채권수익률 | PSO-LSTM, SVM, GRU, VAR, ARIMA/GARCH | PSO-LSTM RMSE 0.0958 |

# Ⅱ. 데이터 및 방법론

1. 종속변수 데이터

본 연구는 2020년 1월부터 2024년 12월까지의 일별 원/달러 환율 종가를 예측 대상으로 한다. 이 시기는 코로나19 팬데믹의 충격, 글로벌 통화정책의 급격한 전환, 지정학적 리스크의 확산 등 서로 다른 시장 국면을 모두 포함하고 있어 환율 변동성이 경제 전반에 미치는 파급효과를 분석하기에 적합하다. 실제로 2020년 이후 원/달러 환율은 팬데믹 초기의 급락, 2021년 저점 국면, 2022년 미국 연방준비제도의 초고속 금리 인상과 한·미 금리차 확대, 정권교체에 따른 국내 정치·정책 변화, 러시아·우크라이나 전쟁과 중동 지역 긴장 등 국내외 요인이 복합적으로 작용하면서 큰 폭의 상승세를 나타냈다.

<그림 1>에 제시된 환율 추이를 살펴보면, 2021년 초 약 1,070원 수준까지 하락한 이후, 2022년 5월에는 1,300원을 돌파하며 환율 변동성이 급격히 확대되었음을 확인할 수 있다. 특히 2021년 상반기의 일평균 환율 등락폭이 ±5원 수준에 불과했던 반면, 2022년 상반기에는 ±15원 수준으로 확대되었으며, 일간 표준편차 기준 환율 변동성 지표 또한 같은 기간 0.4%에서 1.2%로 세 배 가량 급등하였다. 이처럼 변동성이 커진 환율은 2022년 사상 최대의 무역수지 적자를 야기하였고, 원화 약세를 통해 수입물가 상승과 내수 위축, 소비심리 악화로 이어졌다. 반면 수출기업에게는 단기적 가격 경쟁력을 제공하였으나, 환율 급등에 따른 환 헤지 비용은 전년 대비 35% 증가하였고 중소기업을 중심으로 금융비용 부담이 가중되었다.

라인, 그래프, 스크린샷, 텍스트이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**<그림 1> 원/달러 환율 추이**

2. 독립변수 데이터

원/달러 환율을 예측하기 위해 본 연구에서 활용한 독립변수는 기존 환율 예측 연구에서 활용된 주요 요인을 기반으로 하되, 환율 변동의 구조적 요인과 심리적 요인을 동시에 포괄하도록 설계하였다.

선행연구에서는 주로 거시경제 및 금융시장 지표가 독립변수로 사용되었다. Plakandaras et al.(2015)은 원자재, 금속, 주가지수, 금리, 거시경제 변수 등 약 60개 지표를 활용하여 환율을 예측하였으며[5], 임현욱 외(2021)는 IRS 금리, KTB 수익률, 본드 스왑 스프레드 등 금리·채권시장 변수에 집중하였다[6]. Cao et al.(2020)은 유가와 금값 같은 실물자산 가격과 함께 M1·M2, CPI, PPI, 산업생산 등 거시지표를 포함하였고[7], Ali(2025)는 금리, 인플레이션율, GDP 성장률, 외환보유액을 핵심 변수로 설정하였다[8]. Wang et al.(2021)은 USD/CNY 환율뿐 아니라 나스닥, 다우존스, 상하이종합, 항셍 등 글로벌 주가지수를 결합하여 국제적 요인을 반영하였다[9].

최근에는 감성 분석을 통한 시장 심리 변수화 시도가 활발히 이루어졌다. Mohan et al.(2019)은 국제 뉴스 기사에서 추출한 감정 점수를 S&P500 주가와 결합하였으며[10], Jing et al.(2021)은 투자자 게시글의 감정을 CNN 기반 분류기로 산출해 주가 예측에 반영하였다[11]. 정가연 외(2024)는 KLUE-BERT 기반 뉴스 감정 점수를 거시 변수와 함께 투입하여 예측 정확도를 높였고[12], Ding et al.(2025)은 환율 관련 뉴스·댓글 감정을 교차환율, 원자재, 지수, 채권수익률과 결합하여 환율 예측 성능을 향상시켰다[13].

이러한 선행연구의 검토를 통해, 본 연구는 기존의 주요 독립변수들을 포괄하면서도 그 범위를 확장하였다. 우선 종속변수로는 원/달러 환율을 설정하였으며, 교차환율로는 USD/JPY와 USD/CNY를 포함하였다. 이는 원화 환율과 경쟁적 관계에 있는 주요 통화의 변동성을 반영하기 위함이다. 또한 국내외 주식시장 지표인 KOSPI, KOSDAQ, 다우존스, S&P500을 포함하여 자본 흐름과 위험 선호도의 영향을 고려하였다. 원자재 변수로는 유가(WTI), 금, 구리, 니켈, 알루미늄을 채택하여 글로벌 경기 사이클과 인플레이션 기대를 반영하였으며, 거시경제 변수로는 한국과 미국의 정책금리, 소비자물가지수, 생산자물가지수, 경상수지, 통화량(M1, M2)을 선정하였다. 이는 환율 결정에 직접적으로 작용하는 금리 차, 물가 수준, 대외거래, 유동성 요인을 반영하기 위함이다. 아울러 리스크 지표로는 VIX, OVX, KSVKOSPI를 포함하여 금융시장의 불확실성과 위험회피 성향을 계량화하였다.

3. 파생변수 데이터

(1) 이벤트 데이터

본 연구에서는 GDELT로부터 추출한 다섯 가지 핵심 이벤트 기반 변수를 활용하였다. 즉, (1) 전체 이벤트 수, (2) 평균 감성 톤, (3) 한국 관련 이벤트 톤, (4) 전체 문서 수, (5) 한국 관련 문서 수이다. 이벤트 지표는 기존의 거시적 구조적 요인(금리, 주가, 원자재 등) 이외에 '사건 기반 리스크 및 심리적 충격'을 반영할 수 있는 변수로서, 본 연구가 지닌 독창성을 강화해 준다.

(2) 감성분석 데이터

본 연구의 뉴스 감성 분석을 위한 데이터는 네이버 금융 뉴스의 ‘환율’ 섹션에서 2020년 1월 1일부터 2024년 12월 31일까지 수집하였다. 기사 제목과 본문에서 “원화” 또는 “원/달러” 키워드가 포함된 경우만 선별하였다. 이후 텍스트 정제 과정에서 대괄호와 소괄호 안의 불필요한 문구, 이메일 주소, 광고성 표현, “=” 기호 이전의 내용을 정규표현식을 활용하여 제거하였으며, 동일한 제목과 본문을 가진 기사는 중복 수집된 동일 뉴스로 간주하여 하나만 남겼다. 또한 내용이 비어 있거나 공백만 포함된 기사, 그리고 분석 목적과 무관한 기사는 정보 노이즈를 최소화하기 위해 제외하였다.

뉴스 본문은 길이가 다양하고 불필요한 수식어, 문맥적 중복, 장황한 서술을 포함하는 경우가 많아, 이를 그대로 감성 분석에 활용할 경우 모델의 입력 효율성과 분석 정확도에 제약이 발생할 수 있다. 따라서 본 연구에서는 감성 분석 전처리 과정의 일환으로 뉴스 본문 요약을 수행하였다. 요약 모델은 HuggingFace Hub에서 제공되는 한국어 요약 모델 중 다운로드 수가 높은 mT5-multilingual-XLSum, t5-base-korean-summarization, KoBART-summary-v3 세 가지를 비교하였다. 성능 평가는 AI Hub 한국어 문서 요약(신문기사) 데이터셋을 사용하였으며, 지표는 ROUGE-1, ROUGE-2, ROUGE-L을 적용하였다.

ROUGE는 “Recall-Oriented Understudy of Gisting Evaluation”의 약자로, 자동으로 생성된 텍스트를 평가하기 위한 지표 집합을 의미한다. 보통 텍스트 요약(TS, Text Summarization) 알고리즘이 생성한 요약의 품질을 평가하는 데 사용된다[22]. ROUGE-1은 단어 단위의 중복 정도를 나타내며, 요약이 원문 핵심 어휘를 얼마나 잘 포함하는지 평가한다. ROUGE-2는 연속된 2-gram 중복을 기준으로 하여, 문맥적 연결성과 유창성을 간접적으로 반영한다. ROUGE-L은 가장 긴 공통 부분수열(Longest Common Subsequence, LCS)을 활용하여, 요약이 원문의 전체 문장 구조와 맥락을 얼마나 잘 보존하는지를 평가한다[26]. (ADD)

그 결과 <표 3>에서 보듯 KoBART-summary-v3가 모든 지표에서 가장 높은 성능을 기록하여 최종 요약 모델로 채택하였다.

**<표 3 > 요약 모델 성능 표**

|  |  |  |  |
| --- | --- | --- | --- |
| 모델 | ROUGE-1 | ROUGE-2 | ROUGE-L |
| mT5\_XLSum | 0.2683 | 0.0762 | 0.2613 |
| T5-korean-summarization | 0.2356 | 0.0604 | 0.2322 |
| KoBART-summary-v3 | **0.4128** | **0.1632** | **0.3835** |

감성 분석은 GPT-4o-mini 기반 LLM 프롬프트 분류 방식을 채택하였다. 기사별로 제목, 본문, 요약문을 각각 입력으로 사용하고, GPT-4o-mini가 반환한 라벨을 수집하였다. 라벨 체계는 금융 맥락을 반영하여 긍정(원/달러 하락·원화 강세·위험선호 확대·완화적 환경), 부정(환율 상승·원화 약세·위험회피 확대·긴축적 환경), 중립(방향성 판단 곤란)으로 정의하였다. GPT-4o-mini에는 few-shot 프롬프트 튜닝을 적용하여 별도 파인튜닝 없이도 금융 맥락에 특화된 감성 분류를 수행하였다.

라벨 결과는 수치화하여 긍정=+1, 중립=0, 부정=-1로 변환 후 날짜별 평균 점수를 산출해 일별 감정 지표를 구성하였다. 또한 분석 정밀도를 높이기 위해 분산 값과 긍·중립·부정 비율 변수도 함께 도출하였다.

4. 데이터 전처리 및 통계량

결과적으로 본 연구는 거시·금융·원자재·리스크 변수와 더불어 이벤트 기반 변수 및 뉴스·댓글 감성 지표까지 포괄하여, 환율 변동의 구조적 요인과 단기 시장 심리 요인을 동시 고려하는 통합적 예측 프레임워크를 구축하고자 한다. <표 2>

**<표 2> 환율 예측에 사용한 데이터**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 대분류 | 출처 | 데이터명 | 기간 | 단위 | 추출 변수 | 활용 변수 |
| **목표 변수** | Yahoo Finance | KRW/USD | 2020.1–2024.12 | 일 | 종가 | Target (원/달러 환율) |
| **교차 환율** | Yahoo Finance | USD/JPY, USD/CNY | 2020.1–2024.12 | 일 | 종가 | USD/JPY 종가, USD/CNY 종가 |
| **시장 지표** | Investing.com | KOSPI, KOSDAQ, 다우존스, S&P500 | 2020.1–2024.12 | 일 | 종가, 거래량 | KOSPI 종가·거래량, KOSDAQ 종가·거래량, 다우존스 종가, S&P500 종가 |
| **원자재** | Yahoo Finance, Investing.com | WTI, 금(XAU), 구리, 니켈, 알루미늄 | 2020.1–2024.12 | 일 | 종가 | WTI 종가, 금, 구리, 알루미늄, 니켈 종가 |
| **거시 지표** | ECOS, FRED, 통계청, 한국은행 | 정책금리, CPI, PPI, 경상수지, 통화량(M1, M2) | 2020.1–2024.12 | 월 | 한국/미국 정책금리, 소비자물가지수, 생산자물가지수, 경상수지, M1, M2 | 동일 |
| **리스크 지표** | Yahoo Finance | VIX, MOVE, OVX, KSVKOSPI | 2020.1–2024.12 | 일 | 주식·채권·유가 변동성, 코스피 불안 지수 | 동일 |
| **이벤트 지표** | GDELT | events\_cnt, tone\_mean, tone\_kor\_mean, gkg\_doc\_cnt, gkg\_kr\_loc\_cnt | 2020.1–2024.12 | 일 | 이벤트 수, 평균 톤, 한국 관련 톤, 전체 GKG 문서 수, 한국 관련 문서 수 | 동일 |
| **감정 지표** | Naver News, | 환율 관련 뉴스 | 2020.1–2024.12 | 일 | 뉴스 감정 점수 | 감정 평균, 분산, 비율 |

위 데이터들의 결합과정에서 수행했던 전처리 블라블라~~~

데이터는 학습(80%), 테스트(20%)로 분할하였으며, 모든 수치형 변수는 RobustScaler로 정규화하였다. RobustScaler는 중앙값과 사분위 범위를 활용하여 극단값(outlier) 영향을 최소화하는 장점이 있다.

변수들의 통계량 (N, Mean, Median, Std. Q1, Q3)

4. 분석 방법론: LSTM, GRU, CNN-LSTM, CNN-GRU

환율 예측 모델은 LSTM과 GRU 및 하이브리드 모델(CNN-LSTM, CNN\_GRU)을 대상으로 성능을 비교하였다.

최근 외환시장의 복잡성과 비정형적 거시경제 요인을 반영한 시계열 예측 분야에서는 통계 모형을 넘어 딥러닝 기반 아키텍처가 각광받고 있다. 그 중에서도 순환신경망(RNN) 계열의 LSTM(Long Short-Term Memory)과 GRU(Gated Recurrent Unit)는 장기 의존성 학습과 계산 효율성 면에서 주목할 만한 성능을 보인다.

LSTM은 Hochreiter와 Schmidhuber(1997)가 제안한 구조로, 메모리 셀과 게이트 메커니즘을 통해 vanishing gradient 문제를 완화하며 긴 시퀀스 학습에 강점을 가진다[28]. 반면 GRU는 update gate와 reset gate 두 가지 게이트만을 사용하여 구조가 단순하면서도 LSTM에 근접한 성능을 제공하며, 연산 효율성이 뛰어나다는 장점이 있다[29].

외환 예측에서도 이들 모델의 유용성이 다수 연구에서 입증되었다. 예를 들어 Qu와 Zhao(2019)는 EUR/USD 환율 예측에서 LSTM이 기존 RNN보다 RMSE와 MAE가 낮고, 예측 정확도 또한 우수함을 보고하였다[30]. Agustin과 De Melin(2024)은 ARS/USD 예측에서 GRU가 CNN보다 모든 지표에서 더 나은 성능을 보였으며, 이는 환율 데이터의 시계열적 패턴 포착에 순환 구조가 효과적임을 시사한다[31]. 또한 Islam & Hossain(2021)은 EUR/USD 등 주요 통화쌍을 대상으로 GRU와 LSTM을 결합한 하이브리드 모델을 제안하였고, 단일 모델 대비 예측 정확도가 향상됨을 확인하였다[32]. 종합하면, 선행연구들은 LSTM과 GRU가 외환 시장 예측에서 높은 잠재력을 지니며, 상황에 따라 하이브리드 접근이 예측 성능 개선에 유의미하게 작동할 수 있음을 보여준다.

딥러닝 기반 시계열 예측은 단순한 트렌드 분석을 넘어, 변수 간 복잡한 상호작용을 동시에 포착할 수 있는 하이브리드 구조로 발전하고 있다. 이 가운데 CNN-LSTM과 CNN-GRU는 Convolutional Neural Network(CNN)의 국소 특징 추출 능력과 순환신경망(RNN) 계열의 시계열 패턴 학습 능력을 결합한 대표적 모델이다. CNN-LSTM은 입력 시계열 데이터에 1차원 합성곱 계층(1D Convolution Layer)을 적용해 의미 있는 지역적 특징(feature)을 먼저 추출한 뒤, 이를 LSTM에 전달하여 장기적 의존성을 학습한다. CNN-GRU는 동일한 원리를 따르되, GRU의 단순화된 게이트 구조를 통해 연산 효율성과 학습 속도를 강화한 형태이다.

금융시장 예측에서도 이러한 하이브리드 모델은 단일 구조보다 우수한 성능을 보여왔다. 예컨대 Lu et al.(2020)은 주가 예측 실험에서 CNN-LSTM 모델이 가장 높은 예측 정확도를 기록했다고 보고하였으며[33], Jing et al.(2021) 또한 CNN-LSTM을 활용해 투자자 심리를 반영한 주가 예측에서 기존 모델을 능가하는 성과를 입증하였다[11].

따라서 본 연구는 뉴스 요약 및 댓글 감정 점수와 같은 비정형 텍스트 기반 심리 변수를 포함한 시계열 데이터를 CNN-LSTM과 CNN-GRU 모델에 적용하였다. 이로써 CNN이 시장 심리의 국소적 변동을 효과적으로 포착하고, RNN 계열이 이를 시간 축에서 장기적 흐름으로 연결함으로써, 원/달러 환율 예측의 정밀도를 한층 높일 수 있을 것으로 기대된다.

5. 예측 성능평가 지표

본 연구에서는 원/달러 환율 예측 모델의 성능을 평가하기 위해 MSE(Mean Squared Error)와 MAPE(Mean Absolute Percentage Error) 두 가지 회귀 지표를 활용하였다. 두 지표 모두 실제값과 예측값 간의 차이를 기반으로 하며, 값이 낮을수록 모델의 예측 성능이 우수함을 의미한다.

MSE는 예측 오차의 제곱 평균으로, 개별 오차를 제곱하여 평균하기 때문에 상대적으로 큰 오차에 더 큰 가중치를 부여한다. 따라서 모델이 극단적인 오차를 얼마나 잘 억제하는지를 평가하는 데 적합하다. 반면, MAPE는 실제값 대비 예측 오차의 상대적 크기를 백분율로 계산한 값으로, 모델 성능을 직관적으로 해석할 수 있다는 장점이 있다.

두 지표 모두 예측값과 실제값​의 차이를 기반으로 하며, 수치가 낮을수록 모델의 환율 예측력이 높다고 판단할 수 있다. MSE는 모델의 전반적인 안정성과 큰 오차 억제 능력을 확인하는 데 적합하며, MAPE는 실제 비즈니스·금융 현장에서 오차를 쉽게 해석하고 활용할 수 있다는 장점이 있어, 두 지표를 함께 활용함으로써 모델 성능을 다각도로 평가하였다.

6. 최종 방법론 선별을 위한 하이퍼파라미터 최적화

딥러닝 모델의 학습 성능은 입력 데이터의 분포에 크게 의존하므로, 적절한 스케일링 기법의 선택은 예측 안정성과 성능 확보에 핵심적이다. 본 연구에서는 실제 분석에 앞서 여러 스케일링 기법의 효과를 검증하였다. 동일한 입력 변수 조건(Lookback=5)과 LSTM 단일 모델을 사용하였으며, 3개 시드(42, 55, 68)를 고정해 반복 학습한 후 평균과 표준편차를 집계하였다. 비교된 스케일링 방법은 StandardScaler, MinMaxScaler, RobustScaler, 비스케일링(원자료) 네 가지이며, 평가 지표로는 MSE, RMSE, MAE, MAPE를 적용하였다. 실험 결과, RobustScaler 적용 시 가장 낮은 RMSE(7.59), MAE(5.74), MAPE(0.417)를 기록하며 전반적으로 우수한 성능을 보였다. 특히 표준편차 또한 가장 낮아 결과의 안정성이 확보되었다. 반면 StandardScaler는 RMSE와 MAE가 가장 높고 변동성도 커 상대적으로 불리하였다. 이는 환율 데이터가 아웃라이어나 비대칭적 분포를 가질 가능성이 크며, 이에 강건한 RobustScaler가 가장 적합한 전처리 방법임을 시사한다. 따라서 이후의 모든 분석에서는 RobustScaler를 표준 전처리 기법으로 채택하였다.

**<표 5> 스케일링 기법별 성능 비교**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Scaler** | **MSE (평균±표준편차)** | **RMSE (평균±표준편차)** | **MAE (평균±표준편차)** | **MAPE (%) (평균±표준편차)** |
| **Robust** | 57.61 ± 0.87 | **7.59 ± 0.06** | **5.74 ± 0.06** | **0.417 ± 0.004** |
| MinMax | 58.67 ± 0.52 | 7.66 ± 0.03 | 5.80 ± 0.04 | 0.421 ± 0.002 |
| None | 58.66 ± 0.77 | 7.66 ± 0.05 | 5.83 ± 0.06 | 0.424 ± 0.004 |
| Standard | 59.54 ± 1.90 | 7.72 ± 0.12 | 5.89 ± 0.16 | 0.427 ± 0.011 |

학습 파라미터는 epoch 80, batch size 64, Adam 옵티마이저, MSE 손실 함수로 통일하였으며, EarlyStopping(patience=6)과 ReduceLROnPlateau(factor=0.5, patience=3, min\_lr=1e-5)를 적용해 과적합을 방지하였다. 모든 실험은 SEED=42로 고정해 재현성을 확보하였다. <표 6>은 모델별 세부 하이퍼파라미터를 요약한 것이다.

**<표 6> 모델별 세부 하이퍼파라미터**

|  |  |  |
| --- | --- | --- |
| 알고리즘 | 주요 파라미터 | 값 |
| **LSTM** | Unit Numbers | 64 → 32 (2-스택) |
| Dropout Ratio | 0.1 |
| Dense Units | 1 |
| **GRU** | Unit Numbers | 64 → 32 (2-스택) |
| Dropout Ratio | 0.1 |
| Dense Units | 1 |
| **CNN-LSTM** | Filters | 32 |
| Kernel Size | 3 (causal padding) |
| Activation | ReLU |
| Batch Normalization | 적용 |
| LSTM Units | 64 → 32 |
| Dropout Ratio | 0.1 |
| Dense Units | 1 |
| **CNN-GRU** | Filters | 32 |
| Kernel Size | 3 (causal padding) |
| Activation | ReLU |
| Batch Normalization | 적용 |
| GRU Units | 64 → 32 |
| Dropout Ratio | 0.1 |
| Dense Units | 1 |

Ⅲ. 연구결과

3.1 예측 성능평가

본 연구는 원/달러 환율 예측을 위해 LSTM, GRU, CNN-LSTM, CNN-GRU 네 가지 딥러닝 아키텍처를 적용하였다. 데이터셋은 시계열 기반으로 80%를 학습용, 20%를 테스트용으로 분할하였으며, 학습 과정에서 발생할 수 있는 불확실성을 완화하기 위해 시드(seed) 42, 55, 68을 적용하여 세 차례 반복 학습을 수행하였다. 최종 성능은 이들의 평균값으로 평가하였으며, 모든 실험에서 하이퍼파라미터는 batch size=64, epoch=80, 최적화 알고리즘은 Adam, 손실 함수는 MSE로 고정하였다. 입력 시퀀스의 길이는 lookback window를 5, 10, 20, 30, 60, 90일로 설정하여 과거 기간 길이에 따른 영향을 분석하였고, 예측 목표는 다음 날 환율로 설정하였다.

피처 조합은 네 가지 케이스로 구분하였다. (1) 거시·시장 변수만을 포함한 Baseline, (2) Baseline에 GDELT 이벤트 변수를 추가한 조합, (3) Baseline에 뉴스 요약 감성 지표를 추가한 조합, (4) Baseline에 이벤트 지표와 감성 지표를 모두 추가한 통합 조합이다.

<표 7>은 모델별 MSE 기준 상위 20개의 성능 지표를 제시한 결과이다. 이벤트·감성 통합 입력과 CNN-LSTM 아키텍처, 그리고 lookback=20의 조합이 가장 낮은 오차를 달성하였다. 해당 구성의 교차-시드 평균 성능은 MSE 47.71, MAPE 0.39%로 나타났다. 이는 사건·감성 기반 신호가 환율 예측에 실질적인 설명력을 제공함을 시사한다.

<표 8>는 케이스와 lookback 조합별로 네 모델과 세 시드의 성능을 평균으로 계산한 결과를 보여준다. 특히 lookback=20에서 이벤트·감성 통합(49.94)이 베이스라인(50.51), 이벤트 단독(52.76), 감성 단독(54.13)을 모두 상회하여, 거시·시장 변수에 사건 강도·톤 및 뉴스 감성 신호를 결합할 때 예측력이 유의하게 개선됨을 확인하였다. 반면 lookback이 60~90일로 길어질수록 오차가 증가하는 경향이 나타났는데, 이는 본 데이터셋에서 약 20영업일 규모의 중기 윈도우가 정보 효율성이 가장 높으며, 과도한 윈도우 확장은 신호 희석 및 비정상 구간 혼입으로 이어질 수 있음을 의미한다.

또한 <표 7>에서 CNN-LSTM과 CNN-GRU가 상위권을 다수 차지하여, CNN을 통한 국소 패턴 추출과 RNN 기반 시퀀스 학습을 결합한 하이브리드 구조가 환율 시계열의 단기적 구조적 변동과 비선형 상호작용을 포착하는 데 효과적임을 보여준다. 이러한 결과는 사건·감성 정보가 환율의 단기 방향성과 변동성 예측에 실질적인 기여를 한다는 점, 그리고 최적의 시간창 선택 및 하이브리드 아키텍처 설계가 실무적 예측력 향상에 핵심적이라는 점을 동시에 뒷받침한다.

**<표 7> 모델별 MSE 기준 상위 20개 성능 지표**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **case** | **lookback** | **model** | **MSE\_mean** | **MAPE\_mean** |
| Macro+Event+Sentiment | 20 | CNN\_LSTM | 47.711801 | 0.390860 |
| Macro-Only (Baseline) | 20 | CNN\_GRU | 48.835108 | 0.394033 |
| Macro+Event | 30 | CNN\_LSTM | 49.254826 | 0.395459 |
| Macro+Sentiment | 20 | LSTM | 49.451321 | 0.396511 |
| Macro-Only (Baseline) | 20 | CNN\_LSTM | 49.452279 | 0.397167 |
| Macro+Event+Sentiment | 20 | GRU | 49.751230 | 0.403096 |
| Macro+Sentiment | 20 | CNN\_LSTM | 49.945315 | 0.399908 |
| Macro+Event+Sentiment | 20 | LSTM | 50.127059 | 0.401966 |
| Macro+Event+Sentiment | 10 | CNN\_LSTM | 50.505217 | 0.400111 |
| Macro-Only (Baseline) | 20 | GRU | 50.550673 | 0.406172 |
| Macro+Event | 20 | GRU | 50.557137 | 0.404023 |
| Macro+Event+Sentiment | 30 | CNN\_LSTM | 50.782956 | 0.401163 |
| Macro+Event+Sentiment | 10 | LSTM | 50.798627 | 0.399336 |
| Macro+Event | 5 | CNN\_LSTM | 51.021904 | 0.398491 |
| Macro+Event | 30 | LSTM | 51.056315 | 0.399714 |
| Macro-Only (Baseline) | 10 | CNN\_LSTM | 51.067476 | 0.402199 |
| Macro-Only (Baseline) | 10 | LSTM | 51.123407 | 0.403829 |
| Macro+Event | 5 | CNN\_GRU | 51.304119 | 0.400137 |
| Macro+Event | 10 | CNN\_LSTM | 51.339266 | 0.404624 |
| Macro+Event | 10 | LSTM | 51.402042 | 0.401379 |

**<표 8> Case/lookback 조합별 성능**

|  |  |  |  |
| --- | --- | --- | --- |
| **case** | **lookback** | **MSE\_mean** | **MAPE\_mean** |
| Macro+Event+Sentiment | 20 | 49.940599 | 0.401737 |
| Macro-Only (Baseline) | 20 | 50.514034 | 0.401960 |
| Macro+Event | 10 | 52.464776 | 0.406626 |
| Macro+Event | 5 | 52.479801 | 0.405694 |
| Macro+Sentiment | 10 | 52.598089 | 0.406816 |
| Macro+Event | 20 | 52.765744 | 0.410007 |
| Macro+Event+Sentiment | 10 | 52.786881 | 0.406674 |
| Macro+Event | 30 | 53.521021 | 0.412283 |
| Macro+Event+Sentiment | 5 | 53.590648 | 0.408921 |
| Macro+Event+Sentiment | 30 | 53.725530 | 0.411322 |
| Include\_Summary | 20 | 54.132825 | 0.415138 |
| Macro-Only (Baseline) | 10 | 54.220277 | 0.415916 |
| Macro-Only (Baseline) | 5 | 54.833656 | 0.414256 |
| Macro+Sentiment | 30 | 55.538956 | 0.418059 |
| Macro+Event+Sentiment | 90 | 55.680280 | 0.414756 |
| Macro+Event+Sentiment | 60 | 55.707969 | 0.417107 |
| Macro-Only (Baseline) | 90 | 55.939171 | 0.416460 |
| Macro+Sentiment | 5 | 56.672678 | 0.422018 |
| Macro-Only (Baseline) | 30 | 56.829324 | 0.425955 |
| Macro+Event | 60 | 57.186244 | 0.421691 |
| Macro-Only (Baseline) | 60 | 57.362518 | 0.422097 |
| Macro+Event | 90 | 58.036334 | 0.423962 |

3.2 변수 중요도 분석

모델이 어떤 정보를 통해 예측력을 확보했는지를 정량화하기 위해 Permutation Importance를 적용하였다. Permutation Importance는 특정 피처의 값을 무작위로 섞어 모델 예측 성능(MSE)의 변화를 측정하는 방식으로, 모델에 내재된 비선형성과 상호작용 구조를 고려하면서도 직관적인 변수 기여도를 산출할 수 있다. 이러한 방식은 회귀계수에 직접 의존하지 않기 때문에 딥러닝·트리 기반 모델 등 ‘블랙박스’ 모델의 해석 가능성을 높이는 데 널리 활용된다.

<그림 2>는 본 연구에서 가장 우수한 성능을 보인 조합(Macro+Event+Sentiment, CNN-LSTM, lookback=20)으로 측정한 결과이다. 분석에 따르면 Macro/Market 변수군의 비중이 압도적으로 컸다. 구체적으로는 금 가격(XAU), 알루미늄·구리 등 원자재 가격, 미국 정책금리 및 미 국채 10년물 수익률(us10y), OVX와 같은 변동성 지표, KOSPI/KOSDAQ 가격·거래량이 상위권에 포진하였다. 이는 글로벌 금리·상품·주식시장의 공분산 구조가 원/달러 환율 변동에 핵심 설명력을 갖는다는 점을 재확인시켜 준다.

또한 뉴스 감성 요약 지표의 분산(score\_summary\_var)이 높은 중요도를 보였다. 이는 단순한 긍·부정 수준보다 뉴스 톤의 불확실성(변동성) 자체가 단기 환율 예측에 중요한 신호임을 의미한다. 반면 이벤트(GDELT) 변수군은 events\_cnt, gkg\_doc\_cnt, events\_tone\_kor\_mean, gkg\_kr\_loc\_cnt 등 일부 변수가 기여했으나, 그룹 합계 기준으로는 Macro/Market 대비 상대적으로 낮은 중요도를 나타냈다. 이러한 차이는 이벤트 데이터의 잡음, 보도 편향, 또는 사건 영향의 단기·국면 의존성에서 기인할 수 있다.

텍스트, 스크린샷, 소프트웨어, 컴퓨터 아이콘이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**<그림 2> Permutation Importance (Top 20)**

3.3 시각화

본 절에서는 최적 구성(Macro+Event+Sentiment, CNN-LSTM, lookback=20)을 이용하여 2020년부터 2024년까지의 원/달러 환율 예측 결과를 시각화하였다<그림 3>. 그래프는 실제 값(Actual)과 모델의 예측 값(Predicted)을 전체 구간에 걸쳐 비교하였으며, 학습·검증에 사용된 구간과 테스트 구간은 파란색 점선으로 구분하였다.

전체 시계열 비교 결과, 모델은 환율의 장기적 추세뿐만 아니라 중단기적 변동 패턴을 비교적 안정적으로 추적하는 모습을 보였다. 특히 코로나19 이후 글로벌 유동성 확대 국면(2020–2021), 미 연준의 급격한 긴축과 환율 상승(2022), 그리고 이후 조정 국면(2023–2024)에 걸쳐 방향성과 국면 전환을 대체로 잘 포착하였다. 이는 이벤트·감성 정보가 단순 시계열 신호를 넘어, 국면 전환기에서 유의미한 보조 지표로 작용했음을 시사한다.

또한 <그림 4>는 테스트 구간을 확대하여 실제 값과 예측 값의 오차를 상세히 비교한 것이다. 전반적으로 실제 값의 국지적 급등락을 예측하기에는 한계가 있었으나, 평균 수준과 변동성의 큰 방향성은 안정적으로 재현하였다.

결론적으로, 제안된 최적 모델은 환율의 구조적 추세와 단기적 변동성을 모두 일정 수준 포착하며, 이벤트·뉴스 감성 정보를 추가할 경우 단순 거시변수만을 사용할 때보다 전반적 예측력이 향상됨을 시각적으로 확인할 수 있었다.

그래프, 텍스트, 라인, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**<그림 3> 원/달러 환율 예측 결과**

그래프, 라인, 스크린샷, 도표이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**<그림 4> Test 구간 예측 결과**

Ⅳ. 결론

본 연구는 원/달러 환율 예측을 위해 거시·시장 변수에 이벤트(GDELT) 및 뉴스 감성 요약 지표를 결합하고, LSTM, GRU, CNN-LSTM, CNN-GRU 네 가지 딥러닝 아키텍처를 비교·평가하였다. 학습은 시계열 기반 데이터(2020–2024)를 대상으로 수행되었으며, 학습 과정의 불확실성을 완화하기 위해 세 가지 시드를 적용한 반복 실험으로 평균 성능을 측정하였다.

실험 결과, 이벤트·감성 통합(Macro+Event+Sentiment)과 CNN-LSTM 구조, lookback=20일의 조합이 가장 우수한 예측력을 보였다(MSE=47.71, MAPE=0.39%). 이는 이벤트 및 뉴스 감성 신호가 거시·시장 변수와 결합될 때 환율 예측력이 의미 있게 개선됨을 보여준다. 특히 lookback이 20일일 때 최적 성능을 보였으며, 과도하게 긴 윈도우(60~90일)는 오히려 성능 저하를 초래하는 경향이 나타났다.

변수 중요도 분석에서는 금, 알루미늄, 구리 등 원자재 가격과 미국 정책금리·국채금리, OVX·VIX와 같은 변동성 지표, KOSPI/KOSDAQ 지수 및 거래량 등 거시·시장 변수군이 핵심 설명력을 가지는 것으로 나타났다. 동시에 뉴스 감성 지표의 분산(tonal variability)이 상위권에 위치하여, 단순 긍·부정 평균보다는 뉴스 톤의 불확실성 자체가 단기 환율 변동의 신호로 작용할 수 있음이 확인되었다. 반면 이벤트(GDELT) 변수군은 일부 개별 지표가 기여하였으나, 그룹 단위의 중요도는 상대적으로 낮았다. 이는 이벤트 데이터의 노이즈 및 국면 의존성에 기인할 수 있다.

시각화 결과 또한 제안 모델이 2020–2024 전 기간 동안 환율의 구조적 추세와 단기 변동성을 안정적으로 추적함을 보여주었다. 특히 코로나19 이후 급격한 환율 변동, 미 연준의 긴축 정책기 등 주요 국면에서 방향성을 대체로 잘 포착하였다. 이는 환율 예측에서 사건·감성 정보가 단기적 방향성과 변동성 추정에 보조지표로서 유용함을 시사한다.

종합하면, 본 연구는 환율 예측에 있어 이벤트·감성 정보의 결합 효과, 최적의 시간창 선택의 중요성, CNN-LSTM 기반 하이브리드 구조의 유효성을 실증적으로 제시하였다. 이러한 결과는 금융기관 및 기업의 환위험 관리, 단기 환헤지 전략 수립 등에 실질적 함의를 제공할 수 있다. 다만 이벤트 데이터의 노이즈와 다중공선성 문제로 인한 변수 중요도 해석 한계, 그리고 급격한 외생 충격(예: 지정학적 리스크)에 대한 대응 한계는 향후 보완이 필요하다.

향후 연구에서는 GDELT 외에도 다양한 뉴스·SNS 기반 데이터 소스를 통합하고, 다양한 하이브리드 딥러닝 구조 실험을 통해 사건·감성 정보의 활용 가능성을 한층 강화할 수 있을 것이다.

부록

*A. 변수 상관관계 히트맵*

<그림 5>는 변수 간 상관관계 히트맵을 제시한 것이다. 결과를 보면 환율과 주가, 금리, 원자재 가격 등 주요 거시·시장 지표들은 서로 강한 양(+)의 상관 혹은 음(–)의 상관을 보이며, 전형적인 금융 변수들 간 공분산 구조가 나타난다. 뉴스 감성 지표들 역시 서로 밀접한 상관관계를 형성하고 있음을 확인할 수 있다.

패턴, 스크린샷, 텍스트, 사각형이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**<그림 5> 변수 상관관계 히트맵**

# 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

[1] 한국은행. *환율의 정의*. 한국은행. https://www.bok.or.kr/portal/main/contents.do?menuNo=200407

[2] 김인준, & 이영섭. (2019). *국제경제론*. 박영사.

[3] 한국경제. (2023, 10월 22일). [다산칼럼] 환율 예측이 늘 빗나가는 까닭. *한국경제신문*. https://www.hankyung.com/article/2023102265501

[4] 한국은행. *환율의 변동요인*. 한국은행. https://www.bok.or.kr/portal/main/contents.do?menuNo=200407

[5] Plakandaras, V., Papadimitriou, T., & Gogas, P. (2015). Forecasting daily and monthly exchange rates with machine learning techniques. *Journal of Forecasting*, *34*(7), 560-573.

[6] 임현욱, 정승환, 이희수, & 오경주. (2021). 국고채, 금리 스왑 그리고 통화 스왑 가격에 기반한 외환시장 환율예측 연구: 인공지능 활용의 실증적 증거. *지식경영연구*, *22*(4), 71-85.

[7] Cao, W., Zhu, W., Wang, W., Demazeau, Y., & Zhang, C. (2020). A deep coupled LSTM approach for USD/CNY exchange rate forecasting. *IEEE Intelligent Systems*, *35*(2), 43-53.

[8] Qureshi, A. M. (2025). ML FORECASTING OF EXCHANGE RATES MACHINE LEARNING PREDICTION OF EXCHANGE RATES USING MACROECONOMIC INDICATORS Analysis and Use of Various ML Techniques (SVM, Random Forest). *Available at SSRN 5287273*.

[9] Wang, J., Wang, X., Li, J., & Wang, H. (2021). A prediction model of CNN-TLSTM for USD/CNY exchange rate prediction. *Ieee Access*, *9*, 73346-73354.

[10] Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., & Anastasiu, D. C. (2019). Stock price prediction using news sentiment analysis. 2019 IEEE fifth international conference on big data computing service and applications (BigDataService)

[11] Jing, N., Wu, Z., & Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, *178*, 115019.

[12] 정가연, 이혁제, 이준영, & 이제혁. (2024). 금융 특화 감정분석 모델과 딥러닝 시계열 예측 모델을 활용한 코스피 지수 예측. *대한산업공학회지*, *50*(4), 240-250.

[13] Ding, H., Shi, X., Deng, R., Faroog, S., Dewi, D. A., Abdullah, S. N., & Malek, B. A. (2024). Eur/usd exchange rate forecasting incorporating text mining based on pre-trained language models and deep learning methods. *arXiv preprint arXiv:2411.07560*.

[14] 김우석, & 한규식. (2021). COVID-19 가 원달러환율에 미친 영향. *금융지식연구*, *19*(1), 33-58.

[15] 오인정, & 김우주. (2022). SARIMA 와 ARDL 모형을 활용한 COVID-19 구간별 원/달러 환율 예측. *지능정보연구*, *28*(4), 191-209.

[16] GDELT Project. The GDELT Project. https://www.gdeltproject.org

[17] Schroeder, C., Winterbottom, S., Sitter, R., & Perdigones, J. L. (2013). Big Data Analysis of Human Societal Events: Indications for Forecasting Currency Exchange Rates.

[18] Consoli, S., Tiozzo Pezzoli, L., & Tosetti, E. (2020). Information extraction from the GDELT database to analyse EU sovereign bond markets. Workshop on Mining Data for Financial Applications,

[19] Blanqué, P., Ben Slimane, M., Cherief, A., Le Guenedal, T., Sekine, T., & Stagnol, L. (2022). Monitoring Narratives: an Application to the Equity Market. *Amina and Le Guenedal, Théo and Sekine, Takaya and Stagnol, Lauren, Monitoring Narratives: An Application to the Equity Market (April 7, 2022)*.

[20] Gupta, S., Ranjan, R., & Singh, S. N. (2024). Comprehensive study on sentiment analysis: From rule-based to modern llm based system. *arXiv preprint arXiv:2409.09989*.

[21] Cambria, E., Poria, S., Gelbukh, A., & Thelwall, M. (2018). Sentiment analysis is a big suitcase. *IEEE Intelligent Systems*, *32*(6), 74-80.

[22] Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

[23] Nasiopoulos, D. K., Roumeliotis, K. I., Sakas, D. P., Toudas, K., & Reklitis, P. (2025). Financial Sentiment Analysis and Classification: A Comparative Study of Fine-Tuned Deep Learning Models. *International Journal of Financial Studies*, *13*(2), 75.

[24] Shen, Y., & Zhang, P. K. (2024). Financial sentiment analysis on news and reports using large language models and finbert. 2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS),

[25] Kang, J.-W., & Choi, S.-Y. (2025). Comparative investigation of gpt and finbert’s sentiment analysis performance in news across different sectors. *Electronics*, *14*(6), 1090.

[26] Barbella, M., & Tortora, G. (2022). Rouge metric evaluation for text summarization techniques. *Available at SSRN 4120317*.

[27] Zhang, M., Li, C., Wan, M., Zhang, X., & Zhao, Q. (2024). ROUGE-SEM: Better evaluation of summarization using ROUGE combined with semantics. *Expert Systems with Applications*, *237*, 121364.

[28] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

[29] Zheng, W., & Chen, G. (2021). An accurate GRU-based power time-series prediction approach with selective state updating and stochastic optimization. *IEEE Transactions on Cybernetics*, *52*(12), 13902-13914.

[30] Qu, Y., & Zhao, X. (2019). Application of LSTM neural network in forecasting foreign exchange price. Journal of Physics: Conference Series,

[31] Agustin, F., & De Melin, P. (2024). Comparison of GRU and CNN Methods for Predicting the Exchange Rate of Argentine Peso (ARS) against US Dollar (USD). *International Journal Artificial Intelligent and Informatics*, *2*(1), 9-16.

[32] Islam, M. S., & Hossain, E. (2021). Foreign exchange currency rate prediction using a GRU-LSTM hybrid network. *Soft Computing Letters*, *3*, 100009.

[33] Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN‐LSTM‐based model to forecast stock prices. *Complexity*, *2020*(1), 6622927.