

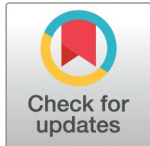
연구 기사

# 데이터 증강, 다변량 기능 및 감정 분석을 통한 정확한 총 소비자 물가지수 예측: 한국 사례 연구

서인재<sup>1</sup>, 김민경<sup>1</sup>, 김종욱<sup>2</sup>, 장백철<sup>1</sup><sup>1</sup> 연세대학교 정보대학원, 서울, 대한민국, <sup>2</sup> 상명대학교 컴퓨터공학과, 서울, 대한민국, 대한민국

이 저자들은 이 작업에 동등하게 기여했습니다.

\*bjang@yonsei.ac.kr (BJ); jkim@smu.ac.kr (JWK)



## 오픈 액세스

**인용:** 서 이, 김 엠, 김 제이 더블유, 장 비 (2025) 데이터 증강, 다변량 특징 및 감정 분석을 통한 정확한 총 소비자 물가 지수 예측: 한국 사례 연구. PLoS One 20(5): e0321530. doi.org/10.1371/journal.pone.0321530

**편집자:** 나말 라트나야케, 도쿄대학교: 도쿄 다이가쿠, 일본

**수신됨:** 접수: 2024년 7월 8일

**수락되었습니다:** 수락: 03월 07, 2025

**게시됨:** 발행일: 2025년 5월 13일

**저작권:** © 2025 Seo 외 이 글은 원저자와 출처를 밝히는 조건으로 모든 매체에서 제한 없이 사용, 배포, 재생산할 수 있는 크리에이티브 커먼즈 저작자표시 라이선스에 따라 배포되는 오픈 액세스 저작물입니다.

**데이터 가용성 정책:** 이 연구에 사용된 데이터 세트는 [https://github.com/mfriendly/prediction\\_KoCPI](https://github.com/mfriendly/prediction_KoCPI)에서 공개적으로 사용할 수 있습니다.

**자금 지원:** 이 연구는 다음 기관의 지원을 받았습니다.

한국연구재단(보조금

number NRF-2024S1A5C3A03046579). 그리고

자금 제공자는 연구 설계, 데이터 수집 및 분석, 출판 결정 또는 원고 작성에 어떠한 역할도 하지 않았습니다.

**경쟁적 이해관계:** 저자는 경쟁적 이해관계가 존재하지 않는다고 선언했습니다.

## 초록

소비자물가지수(CPI)는 전 세계 정책 입안자들이 인플레이션을 모니터링하고 통화 정책 결정을 내리는 데 사용하는 주요 경제 지표입니다. 한국의 경우 CPI는 금리, 재정 정책 기조, 한국은행의 경제 안정 전략에 대한 결정에 큰 영향을 미칩니다. 이러한 중요성을 고려할 때 총 CPI를 정확하게 예측하는 것은 정보에 입각한 의사결정을 내리는 데 매우 중요합니다. 그러나 정확한 예측을 달성하는 데는 몇 가지 어려움이 있습니다. 첫째, 한국의 총 CPI는 다음과 같이 계산됩니다. 12개의 상품 및 서비스 카테고리로 분류된 462개 품목의 가중치 합계입니다. 이러한 이질성으로 인해 소비자 행동과 가격 역학의 모든 변동을 설명하기 어렵습니다. 둘째, CPI 데이터의 월별 빈도는 상대적으로 시계열이 희박하여 분석의 성능을 제한합니다. 또한 정책 변화나 팬데믹과 같은 외부 요인이 CPI에 변동성을 더합니다. 이러한 문제점을 해결하기 위해 크리테오는 네 가지 핵심 요소로 구성된 새로운 프레임워크를 제안합니다: (1) CPI 데이터의 복잡한 패턴을 포착하여 추정 정확도를 향상시키는 하이브리드 컨볼루션 신경망-장단기 메모리 메커니즘, (2) 보다 풍부한 맥락 정보를 위해 보조 변수와 함께 CPI 구성 지수를 통합하는 다변량 입력, (3) 선형 보간을 통한 데이터 증강으로 월별 데이터를 일일 데이터로 변환하여 고도로 파라미터화된 딥러닝 모델에 최적화, (4) 한국 CPI 관련 뉴스 기사에서 도출한 감정 지수로 CPI 변동에 영향을 주는 외부 요인에 대한 인사이트를 제공. 실험 결과, 제안된 모델은 낮은 RMSE 값에서 알 수 있듯이 CPI 예측에서 기존 접근 방식보다 뛰어난 성능을 보였습니다. 이러한 정확도 향상은 보다 시의적절하고 효과적인 경제 정책의 개발을 지원할 수 있는 잠재력을 가지고 있습니다.

## 용어집

약어/용어	정의
BERT	트랜스포머의 양방향 인코더 표현: 자연어 처리 작업을 위해 설계된 딥러닝 모델로, 다음을 지원합니다. 텍스트의 문맥적 이해.
KLUE-BERT	한국어 이해도 평가 BERT: BERT 모델 KLUE 벤치마크의 일부로 한국어 작업에 맞게 미세 조정되었습니다.
KoBERT	트랜스포머의 한국어 양방향 인코더 표현: 한국어에 맞게 사전 학습된 BERT 기반 모델 작업을 이해합니다.
CPI	소비자물가지수: 소비자 및 서비스 바구니의 평균 가격 수준을 나타내는 척도입니다. <a href="#">섹션 1</a> 과 <a href="#">섹션 2</a> 에서는 CPI에 대해 일반적인 용어로 설명합니다. <a href="#">섹션 3</a> 이후부터는 달리 명시되지 않는 한, 이 백서에서는 특히 한국 총 CPI(2015년= 100)를 참조합니다.
CPI 지수	<b>총 CPI:</b> 총 소비자물가지수는 구성 지수의 가중치 합계로 종합 소비자물가지수입니다. <b>구성 지수:</b> 섹터별 가격 변동과 같이 전체 CPI 내의 개별 카테고리 또는 하위 집합을 나타내는 특정 지수입니다. <b>보조 지수:</b> 추가 정보를 제공하는 보조 지수 컨텍스트 또는 기본 CPI의 행동 및 추세에 대한 인사이트를 제공합니다.
감정 색인	에 관련된 한국 뉴스 기사에서 추출한 일일 감정 점수입니다. 별도의 언급이 없는 한 CPI.
FRED	연방 준비 은행 경제 데이터: 공개적으로 액세스 가능한 데이터베이스 연구 및 분석을 위한 경제 및 재무 데이터를 제공합니다.
ANN	인공 신경망: 다음에서 영감을 얻은 머신 러닝 모델 다양한 예측 작업에 사용되는 생물학적 신경망입니다.
CNN	컨볼루션 신경망. 2D 또는 1D CNN을 참조할 수 있으며 1D는 프레임워크에는 CNN이 사용되었습니다.
아리마	자동 회귀 통합 이동 평균: 시계열 예측을 위한 고전적인 통계 모델로, 다음 간의 종속성을 모델링합니다. 관찰.
1D CNN	1차원 컨볼루션 신경망: 시간과 같은 1차원 데이터를 처리하도록 설계된 신경망 아키텍처입니다. 시리즈.
CNN-LSTM	컨볼루션 신경망-장단기 메모리: 특징 추출을 위한 1D CNN과 순차적 종속성을 포착하기 위한 LSTM을 통합한 하이브리드 딥 러닝 모델로, 이를 통해 다변량 시계열 예측의 정확도.
AttentionLSTM	주의 장단기 메모리 네트워크: 주의 메커니즘을 통해 관련 시간 단계의 우선 순위를 지정하는 향상된 LSTM 모델입니다. 순차적 데이터를 사용하여 예측 정확도를 향상시킵니다.
GRU	게이트 리커런트 유닛: 소실 그래데이션을 완화하여 순차적 데이터를 효율적으로 처리하는 순환 신경망 변형 문제.

LSTM	장단기 메모리 네트워크: 종속성을 순차적으로 처리하고 학습하도록 설계된 일종의 순환 신경망 아키텍처입니다. 데이터.
------	---

SVR/SVM	서포트 벡터 회귀/서포트 벡터 머신: SVR은 SVM 원리를 기반으로 하는 회귀 알고리즘으로, 고차원 공간에서 지정된 마진 내에 하이퍼플레인을 맞춰 연속적인 결과를 예측하도록 설계되었습니다. SVM은 주로 다음과 같은 분류 알고리즘입니다. 클래스를 분리할 최적의 하이퍼플레인을 식별합니다.
FFNN	완전 연결(FC) 레이어 2개가 있는 피드포워드 신경망: 시계열 예측을 위한 간단한 신경망 아키텍처입니다. 완전히 연결된 두 개의 순차적 레이어를 사용하여 복잡한 데이터를 캡처합니다. 데이터의 비선형 패턴을 확인합니다.
XGBoost	익스트림 그래디언트 부스팅: 다음에 사용되는 확장 가능한 트리 기반 알고리즘입니다. 시계열 예측을 포함한 회귀 및 분류 작업을 수행합니다.
NLP	자연어 처리: 이해에 중점을 둔 AI 분야입니다, 해석하고 인간의 언어를 생성합니다.
MAE	평균 절대 오차: 이상값에 덜 민감한 지표로 다음을 계산합니다. 잔여물의 평균 크기입니다.
MAPE	평균 절대 백분율 오류: 예측 오류를 정량화하는 지표를 실제 값의 백분율로 표시합니다.
RMSE	평균 제곱근 오차: 모델 평가 지표로 평균 제곱 잔액의 제곱근입니다.
NRMSE	실제 데이터의 표준 편차로 정규화된 RMSE입니다.
SMAPE	대칭 평균 절대 백분율 오류: 오류를 대칭적으로 평가하여 백분율 오류의 과대 추정을 해결하는 메트릭입니다. 를 실제 값과 비교합니다.
TP, FP, TN, FN	<b>정양성(TP):</b> 올바르게 식별된 양성 사례. <b>위양성(FP):</b> 양성으로 잘못 분류된 사례. <b>위음성(TN):</b> 음성으로 올바르게 식별된 사례. <b>거짓 부정(FN):</b> 잘못 분류된 사례.
F1 점수	모델 성능을 평가하기 위한 메트릭으로, 고조파를 나타냅니다. 정밀도 및 회상률의 평균입니다.

## 1 소개

코로나19 팬데믹은 전 세계 경제에 큰 영향을 미쳤으며, 지속적인 가격 상승을 특징으로 하는 급속한 인플레이션으로 이어지고 있습니다. 인플레이션은 구매력을 약화시키고, 화폐 가치를 떨어뜨리며, 일상생활을 불안정하게 만들어 궁극적으로 경제 안정을 저해합니다[1]. 소비자물가지수(CPI)는 인플레이션 추세를 반영하고 정책 입안자와 이해관계자에게 인사이트를 제공하는 중요한 경제 지표 역할을 합니다. 효과적인 경제 계획과 의사결정을 위해서는 정확한 CPI 예측이 필수적입니다. 예를 들어, 한국의 국민연금공단은 CPI 변동에 따라 연금 지급액을 조정하여 다음을 보장합니다.

공정경쟁확인할 수 있으며, 한국은행은 경제 안정을 위해 인플레이션 목표치를 CPI로 설정하고 있습니다[2].

CPI의 근본적인 중요성을 고려할 때 신뢰할 수 있는 예측은 매우 중요합니다. 그럼에도 불구하고 CPI를 예측하는 데는 여러 가지 어려움이 있습니다. 예를 들어 팬데믹과 같은 외부 이벤트는 CPI 역학을 크게 뒤흔들 수 있습니다. 팬데믹 기간 동안 필수품(예: 식료품, 보호 장비)에 대한 수요가 증가하고 여행 관련 제품에 대한 수요가 감소하는 등 소비 패턴이 급격하게 변화했습니다[3]. 공급망 중단

경제 회복은 특히 운송, 식품, 숙박 등의 분야에서 인플레이션 압력을 강화했습니다 [4,5]. 확장적 재정 및 통화 정책의 채택()은 인플레이션 압력에 더욱 기여했습니다[6].

CPI 예측의 복잡성은 각각 다른 역할 관계에 따라 영향을 받는 여러 하위 지수로 구성되어 있기 때문에 발생합니다. 한국의 CPI는 식품, 주택, 교통 등 12개 부문의 462개 대표 품목을 기준으로 매월 산출됩니다. 예를 들어, 주거비 상승이 식료품 가격 하락으로 상쇄되어 전체 CPI가 안정적으로 유지되는 경우와 같이 부문별 변동이 전체 CPI에 가려질 수 있습니다. 따라서 정확한 예측을 위해서는 CPI 구성 요소를 활용하는 것이 필수적입니다.

게다가 일반적으로 매월 측정되는 CPI 데이터의 제한된 빈도로 인해 연간 관측 횟수가 12회에 불과하여 시계열 예측이 더욱 어려워집니다. 팬데믹, 정책 변화, 자연재해와 같은 외부 요인은 이러한 노력을 더욱 복잡하게 만듭니다. 뉴스 기사의 감정 분석과 같은 보조 데이터 소스를 통합하면 CPI 변동을 이해하고 예측하는 데 유용한 맥락을 제공할 수 있습니다.

이러한 문제를 해결하기 위해 다변량 데이터와 감정 분석을 활용하는 하이브리드 컨볼루션 신경망-장단기 메모리(CNN-LSTM) 모델을 제안합니다. 이 모델은 다음과 같은 주요 구성 요소를 도입합니다:

- **다변량 입력:** 구성 요소 및 보조 CPI 지수, 뉴스 기사에서 파생된 심리지수를 통합하여 한국 총 CPI에 대한 부문별 역할 관계와 외부 영향을 포착하는 모델[7].
- **데이터 증강:** 선형 보간은 월별 관찰에서 일일 데이터 포인트를 생성하여 딥러닝 기반 모델에 필수적인 데이터 부족 문제를 해결합니다.
- **감정 분석:** 글로벌 이벤트와 관련된 뉴스 기사에서 파생된 감정 점수에서 추가적인 컨텍스트를 제공하여 외부 충격에 대한 모델의 반응성을 향상시킵니다.
- **CNN-LSTM 하이브리드 아키텍처:** CNN은 다변량 데이터에서 특징을 추출할 수 있으며, LSTM은 순차적 종속성을 포착하여 예측 정확도를 향상시킵니다[8].

광범위한 실험을 통해 다변량 입력, 데이터 증강, 감정 분석, CNN-LSTM 메커니즘을 결합하여 기존 방법보다 뛰어난 예측 성능을 달성한 모델이 입증되었습니다.

2장에서는 CPI 예측에 관한 관련 연구를 검토합니다. 3장에서는 모델의 방법론에 설명하고 4장에서는 실험 설정을 설명합니다. 5장에서는 결과를 제시하고 6장에서는 결론을 내립니다.

## 2 관련 작업

### 2.1 CPI 예측을 위한 기존 및 고급 접근 방식

CPI 예측은 오랫동안 전통적인 통계적 방법에 의존해 왔으며, 가장 널리 사용되는 방법 중 하나는 Autoregressive 통합 이동 평균(ARIMA)입니다. Zhang 등[9]은 에서 거시경제 지표를 예측하는 데 있어 ARIMA의 효과를 입증했으며, Ahmar 등[10]은 인도네시아의 CPI 모델에 적용하여 실제 경제 시나리오에서의 정확성을 보여주었습니다. 마찬가지로 Shinkarenko 등[11]은 ARIMA로 우크라이나의 CPI 동학을 분석하여 소비자 물가 동향을 추적할 수 있는 능력을 보여주었습니다. 예를 들어, Liu[12]는 중국의 CPI를 예측하기 위해 ARIMA와 LSTM 모델을 결합하여 두 방법의 강점을 활용하여 예측 정확도를 향상시킨 하이브리드 접근법도 모색했습니다.

마코프 체인 [13], 서포트 벡터 회귀(SVR) 등의 머신러닝 방법(예: 기계 ) [14]/SVM [15]도 CPI 예측에 사용되었습니다. 그러나 이러한 접근 방식은 일반적으로 단변량 입력에 의존하기 때문에 더 복잡한 관계를 모델링하는 데 한계가 있습니다.

딥러닝 방법의 개발로 인해 예 대한 새로운 가능성 이 생겨났습니다. Moshiri 등[16], Chen 등[17], Szafranek 등[18]의 연구와 같은 초기 연구에서는 신경망 이 종종 전통적인 통계적 방법보다 우수한 성능을 보인 것으로 나타났습니다. Mcadam 등 [19]과 Choudhary 등[20]의 국가 간 분석( )은 인플레이션 예측에 대한 이러한 모델의 장점을 강조했습니다. 나카무라 외[21]는 피드 포워드 신경망(FNN)이 미국 CPI를 예측하는 데 특히 효과적이며, 자동 회귀 기준선을 능가한다는 사실을 발견했습니다. 이를 바탕으로 알모소바 등[22] 자하라 등[7]은 인플레이션 예측을 위해 LSTM 네트워크( rks)를 적용하여 기존 모델보다 더 나은 정확도를 달성했습니다. Zheng 등[23]은 페널티 회귀와 혼합 주파수 데이터 샘플링을 결합하고 미디어 및 인터넷 데이터를 통합하여 예측 성능을 크게 향상시켰으므로 이를 더욱 발전시켰습니다. 최근 연구에서는 더 복잡한 딥러닝 접근 방식을 탐구했습니다. 이와 유사하게 Barkan 등은 계층적 순환 신경망( )을 개발하여 세분화된 CPI 구성 요소를 예측함으로써 기존 모델보다 더 나은 결과를 달성했습니다[24].

전반적으로 이러한 연구는 CPI 예측에서 고급 계산 기술의 중요성이 커지고 있음을 반영합니다. 그럼에도 불구하고 대부분의 방법은 주로 단변량 데이터에 초점을 맞추고 있습니다. CPI의 희박한 빈도 데이터를 직접적으로 다룬 연구는 없었으며, 대부분의 연구는 단순한 모델에 의존했습니다.

## 2.2 CNN 또는 LSTM 모델을 활용한 하이브리드 예측 모델

순환 신경망(RNN) 기반 아키텍처는 숨겨진 계층의 과거 정보를 활용하여 입력과 출력을 순차적으로 처리함으로써 순차 데이터를 효과적으로 학습시킵니다. 그러나 입력 시계열 길이가 길어지면 기존 RNN은 장기적인 종속성으로 인해 과거 정보가 손실되는 문제를 겪게 됩니다. 이 문제를 해결하기 위해 Hochreiter et al.

[25]는 시계열 예측에서 유망한 결과를 보여준 LSTM 모델을 제안했습니다[26,27].

예를 들어, Li 등( ) [28]은 감정 분석을 통합한 경우에도 중국 주식 시장을 예측할 때 LSTM이 SVM 및 선형 회귀 모델보다 우수한 성능을 보였음을 입증했습니다. 그러나 LSTM은 시간적 의존성을 유지하는 데 강점이 있기 때문에 복잡한 변수 간 상호작용을 완벽하게 포착하는 데는 한계가 있습니다.

최근 연구에 따르면 CNN 기반 모델은 예측에서 이러한 한계를 ns 해결할 수 있는 것으로 나타났습니다. 예를 들어, 금융 영역에서 호세인자데 등[29]은 특징 추출을 사용하는 주식 시장 예측을 위한 프레임워크인 CNNpred를 제안했습니다.

및 시장 간 상관관계를 분석하여 기존 방식보다 크게 개선했습니다.

두 접근법의 강점을 활용하기 위해 하이브리드 CNN-LSTM 모델이 제안되었습니다. CNN은 특징 추출에 사용되며( ), LSTM은 시계열 데이터에 내재된 시간적 깊이를 처리합니다( n ). 리비에리스 등[30]과 다른 연구자들[31,32]은 금 가격 예측에서 이 하이브리드 접근법의 효과를 입증했는데, CNN은 주요 특징을 추출하고 LSTM은 미래 가치를 예측합니다. 이러한 연구는 복잡한 상호작용과 순차적 패턴을 포착하기 위해 1차원(1D) CNN과 LSTM을 결합하여 예측 정확도를 향상시킬 수 있는 이점을 강조합니다.

그럼에도 불구하고 이러한 하이브리드 모델 중 일부만이 CPI 예측에 적용되었습니다. 게다가 월별 빈도 데이터를 효과적으로 처리하는 것의 중요성을 인식하지 못했습니다.

## 2.3 재무 예측을 위한 감정 분석

감정 분석은 텍스트 데이터를 활용하여 예측 정확도를 높이는 금융 I 예측의 중추적인 도구가 되었습니다. 예를 들어, Nia 등 [33]은 트위터 감성 분석과 머신러닝을 사용하여 코로나19에 대한 거시경제적 대응을 조사함으로써 전 세계 소득 수준별 고용 및 인플레이션 통제에 차이가 있음을 밝혀냈습니다( ). 마찬가지로 송 등 [34]은 소셜 미디어 정서를 RNN 모델에 통합하면 감염병 사례의 예측 정확도가 향상되어 대중의 여론이 질병 확산에 미치는 질적 영향을 강조하는 것을 입증했습니다.

Shapiro 등 [35]은 어휘 기반 스코어링과 머신러닝을 결합하여 경제 뉴스 감성을 측정하고, 부정과 같은 언어적 뉘앙스를 명시적으로 고려하며, 거시경제 변수 및 소비자 감성에 대한 사전 지식을 가진 감성 시계열 지수를 구축했습니다. 반면, Li 등 [36]은 중국 주식 시장 가격을 예측하기 위해 나이브 베이즈 감성 점수를 LSTM 네트워크 및 서포트 벡터 머신(SVM)과 결합하여 감정 분석을 외부 이벤트와 통합하면 예측 정확도를 크게 높일 수 있음을 보여주었습니다.

이러한 접근법을 바탕으로 De Oliveira Carrosia 외 [8]은 포르투갈어로 된 금융 뉴스에 대한 CNN 기반 분석을 사용하여 브라질 주식 시장에 대한 감성 중심 투자 전략을 소개했습니다. 이들의 연구는 단기 및 장기 시나리오 모두에서 랜덤 워크 및 매수 후 보유와 같은 전통적인 전략보다 우수한 수익성을 보여주었습니다. 마찬가지로 Li 등 [37]은 투자자의 감정을 딥러닝 모델에 통합하여 중국 시장에서 주가 예측을 개선하기 위한 감정 분석의 잠재력을 검증했습니다.

한편, 최근 한국어 자연어 처리(NLP)의 발전으로 한국어 Bidirectional Encoder Representations from Transformers (KoBERT) [38] 및 KLUE-BERT [39] 등 한국어 랜 게이지에 맞게 사전 학습된 언어 모델이 도입되었습니다. KoBERT는 위키백과와 뉴스 기사를 포함한 대규모 한국어 텍스트 말뭉치에 대해 훈련되었습니다. SentencePiece 토큰화기를 활용하여 한국어 NLP 작업에서 다국어 BERT보다 지속적으로 우수한 성능을 발휘합니다[40].

KLUE-BERT[39]는 다양한 NLP 작업을 위해 설계된 또 다른 모델로, 공식 기사, 웹 콘텐츠, 구어체 텍스트 등 다양한 형식으로 구성된 62GB 코퍼스로 훈련되었습니다. 형태소 기반 토큰화와 1억 1천만 개의 파라미터를 갖춘 KLUE-BERT는 KLUE 벤치마크 작업에서 우수한 성능을 입증했습니다.

이러한 발전에도 불구하고 국내 금융 예측에서 감성 분석의 활용은 아직 미흡한 실정입니다. 일반적인 정서 기반 예측에 대한 연구가 집중되어 있는 반면, 한국의 소비자물가지수 예측과 같은 구체적인 적용 사례는 부족합니다. KoBERT, KLUE-BERT와 같은 한국어 NLP 사전 학습 모델은 감정 분석 능력이 크게 향상되었지만, CPI 예측 및 기타 시계열 작업에 대한 잠재력은 아직 충분히 실현되지 않았습니다. 또한 기존 연구에서는 데이터 증강 기술을 통합하거나 CNN-LSTM과 같은 고급 딥러닝 기반 모델을 사용하는 경우가 드뭅니다.

또한, 감성 기반 데이터 증강 기법을 제한적으로 적용하는 것은 이러한 격차를 더욱 강조합니다. 현재의 연구는 주로 전통적인 예측 방법에 의존하고 있으며, 특히 다변량 시나리오에서 모델 성능을 향상시킬 수 있는 증강의 잠재력을 간과하고 있습니다. 이러한 기법을 탐구하면 CPI 예측 및 기타 한국 금융 영역의 시계열 예측 과제와 같은 작업에서 예측 정확도를 크게 향상시킬 수 있습니다. 이러한 한계를 해결함으로써 향후 연구에서는 감정 분석을 고급 예측 프레임워크에 통합할 수 있는 새로운 기회를 열 수 있습니다.

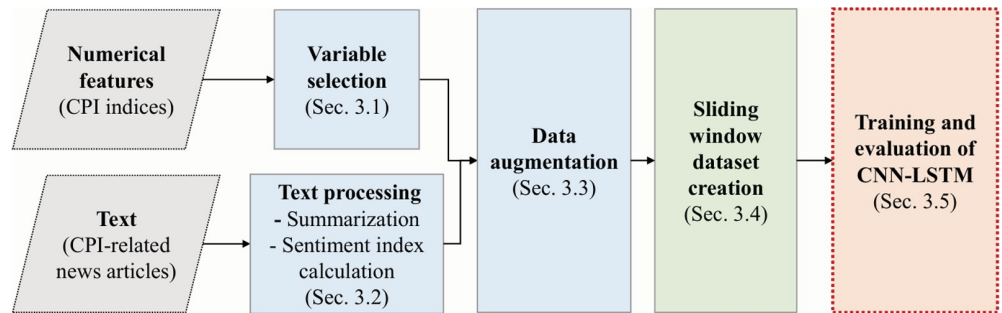


그림 1. CPI 예측을 위한 데이터 전처리 및 모델링 파이프라인.

<https://doi.org/10.1371/journal.pone.0321530.g001>

### 3 방법론

이 연구는 구조화된 전처리 파이프라인을 통해 수치 데이터(CPI 지수)와 텍스트 데이터(CPI 관련 뉴스 아티클)를 통합합니다. 먼저 수치 데이터와 텍스트 데이터 모두에 변수 선택을 적용한 다음(3.1절), 요약 및 감정 지수 계산과 같은 텍스트 처리 기법을 적용합니다(3.2절). 그런 다음 처리된 데이터를 증강하여 모델의 견고성을 높인 다음(3.3절), 순차적 학습을 위한 슬라이딩 윈도우 데이터 세트를 구축합니다(3.4절). 마지막으로, 준비된 데이터 세트는 CNN-LSTM 모델 훈련 및 평가에 사용됩니다(3.5절).

그림 1은 이 파이프라인의 개요로, 기능 선택부터 모델 학습까지의 주요 단계를 보여줍니다. 다음 하위 섹션에서는 각 단계에 대해 자세히 설명합니다.

#### 3.1 CPI 관련 수치 특징 수집 및 선택

CPI 관련 데이터 세트는 연방준비제도 경제 데이터(FRED) API(<https://github.com/mortada/fredapi>)에서 '대한민국 CPI 2015:100'을 검색하여 얻었습니다. 세인트루이스 연방준비은행에서 관리하는 FRED는 국제 통계를 포함한 광범위한 경제 데이터 저장소를 제공합니다.

CPI 관련 지수는 기준 연도를 2015년(2015년= 100)으로 설정하여 상대 지수로 표현합니다. 이 방법론은 지수가 소비자 행동( )과 관련이 있는지 확인합니다.

한국 총 CPI는 물가지수 계산에 널리 사용되는 수정된 라스페이어 공식[41]을 사용하여 계산합니다. 이 방법은 시간에 따른 가격 수준의 변화를 반영하여 기준 기간 대비 현재 기간에 고정된 상품 및 서비스 바구니를 구매하는 데 드는 비용을 계산합니다. 이 공식은 기준 기간과 현재 기간의 개별 품목 가격을 모두 고려하고 기준 기간의 각 소비 점유율에 따라 가중치를 부여하여 계산된 지수가 소비자 지출 패턴과 인플레이션 추세를 정확하게 나타낼 수 있도록 합니다. 라스페이어 공식은 다음과 같이 표현할 수 있습니다:

$$\text{총} = \frac{\sum_{i=1}^n \frac{p_{i,t} q_{i,0}}{p_{i,0} q_{i,0}}}{\sum_{i=1}^n} \times 100 \quad (1)$$

이 공식에서  $p_{i(t)} q_{i(t)}$ 은 현재 기간의 품목  $i$ 의 가격,  $p_{i(t)} q_{i(0)}$ 은 기준 기간의 품목  $i$ 의 가격,  $q_{i(t)} q_{i(0)}$ 은 기준 기간에 소비된 품목  $i$ 의 수량을 나타냅니다. 분모는 기준 기간에  $t$  바구니의 총 비용을 반영하고 분자는 현재 기간의 비용을 나타냅니다[42].



표 1. 한국의 CPI 관련 기능 데이터 세트.

FRED ID	제목	금액 (%)	Coef.	선택됨
<b>기본 색인</b>				
코피알민메이	총 CPI	462 (100.0%)	-	✓
<b>구성 요소 인덱스</b>				
KORCP010000IXOBM	음식 및 무알코올 음료	133 (28.8%)	4.79	✓
KORCP020000IXOBM	, 담배 및 마약류	7 (1.5%)	0.28	
KORCP020000IXOBM	의류 및 신발	30 (6.5%)	0.32	
KORCP040000IXOBM	주택, 수도, 전기, 연료	16 (3.5%)	3.29	✓
KORCP050000IXOBM	가정용품 및 서비스	49 (10.6%)	-	
KORCP060000IXOBM	건강	33 (7.1%)	-	
KORCP070000IXOBM	교통편	32 (6.9%)	1.51	
KORCP080000IXOBM	커뮤니케이션	6 (1.3%)	-	
KORCP090000IXOBM	레크리에이션 및 문화	56 (12.1%)	-	
KORCP100000IXOBM	교육	20 (4.3%)	0.03	
KORCP110000IXOBM	레스토랑 및 호텔	44 (9.5%)	-	
KORCP120000IXOBM	기타 상품 및 서비스	36 (7.8%)	-	
<b>보조 지표</b>				
코르크피코민메이	모든 품목(비식품 비에너지)	-	6.61	✓
코피엔민메이	에너지	-	-	
KORCPGRSE01IXOBM	서비스	-	0.36	
KORCPGRLH02IXOBM	주택이 적은 서비스	-	2.47	✓
KORCPGRHO02IXOBM	주택에 대한 전가 임대료를 제외한 주택	-	-	
KORCP040100IXOBM	실제 주택 임대	-	0.14	
KORCP040400IXOBM	물 공급 및 기타 관련 서비스 주거	-	-	
KORCP040500IXOBM	전기, 가스 및 기타 연료	-	-	
KORCP040300IXOBM	주거지 유지 및 수리	-	-	

<https://doi.org/10.1371/journal.pone.0321530.t001>

한국총생산자물가지수(FRED ID 'KORCP1ALLMINMEI')는 전체 비중의 100%를 차지하는 462개 대표 품목을 포괄하는 종합적인 소비자물가 지표입니다. 주택, 교통, 식료품 등의 품목을 포함하는 구성지수의 가중치 합으로 도출되며, 표 1에 자세히 설명되어 있습니다. 이러한 구성 요소는 소비자 지출에 대한 기여도에 따라 분류됩니다.

구성 요소 지수는 *식료품 및 비알코올 음료*(FRED ID 'KORCP010000IXOBM', 28.8%), *주거, 수도, 전기, 연료*(KORCP040000IXOBM, 3.5%)와 함께 *건강, 교통, 교육* 등의 주요 지출 카테고리를 나타냅니다. 각 구성 요소는 총 소비자 지출에 대한 상대적 기여도에 따라 가중치가 부여됩니다. *모든 항목(비식품 비에너지)*(KORCPICORMINMEI) 및 *주거를 제외한 서비스*(KORCPGRLH02IXOBM)와 같은 보조 지수는 특정 분석 요구에 대한 추가적인 세분성을 제공합니다.

분석에는 22개의 지수가 포함되었습니다: 총 CPI, 12개의 주요 구성 지수, 9개의 보조 지수. 2010년 1월부터 11월까지의 월별 데이터( )를 활용했습니다. 영향력이 낮은 변수를 페널티화하여 유의미한 변수를 식별하는 올가미 회귀[43]를 사용하여 특징 영향력을 평가했습니다. 30% 임계값이 적용되었으며, 0이 아닌 계수는 표 1에 요약되어 있습니다. 올가미 회귀를 사용한 결과, 2개의 구성 지수 및 2개의 보조 지수가 CPI 변동에 중요한 영향을 미치는 것으로 확인되었습니다. 주목할 만한 구성 요소로는 *식료품 및 비알코올 음료*와 *주택, 수도, 전기, 연료*가 있습니다.

그림 2는 0이 아닌 계수 변수의 실제 값을 센티멘트 지수와 함께 보여줍니다. 범례는 올가미 계수를 내림차순으로 변수를 정렬하며, 가장 높은 값을 가진 변수는

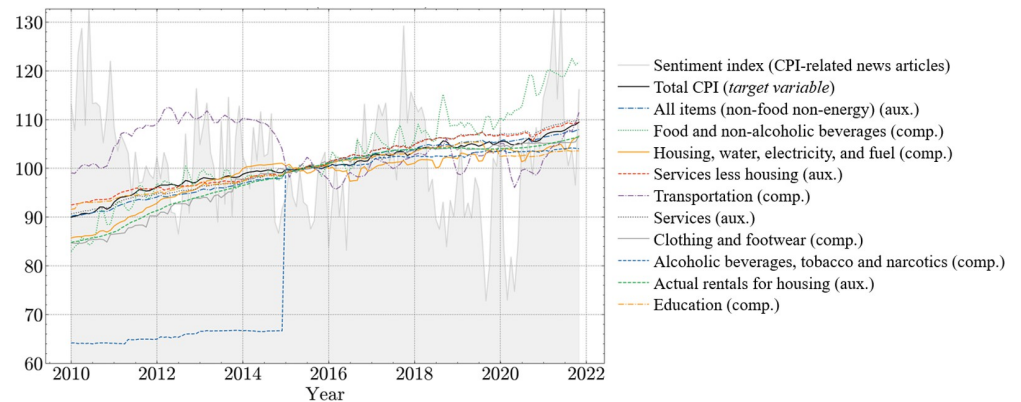


그림 2. 율가미 계수가 0이 아닌 변수의 실제 값과 감정 지수. 범례는 율가미 계수를 내림차순으로 변수를 정렬했으며, 가장 영향력이 큰 것은 총 CPI입니다.

<https://doi.org/10.1371/journal.pone.0321530.g002>

총 CPI 다음의 영향력 있는 변수입니다. 보조 변수와 구성 요소 변수가 혼합되어 있어 구성 요소 비중이 높다고 해서 반드시 더 높은 영향력을 나타내는 것은 아닙니다. 특히 교통과 주류는 뚜렷한 추세를 보이지만 여전히 영향력이 있는 것으로 나타났습니다. 음영 처리된 영역은 감정 지수를 나타냅니다. 해당 율가미 계수 값은 표 1에 나와 있습니다.

특히, *식료품 및 주거*와 같은 구성 지수는 *교육 및 통신*과 같은 범주에 비해 더 큰 변동성을 보이며, 이는 전체 CPI에 미치는 영향력이 더 크다는 것을 반영합니다. 또한 감정 지수는 특히 코로나19 팬데믹과 같은 외부 충격이 발생했을 때 중요한 CPI 변화와도 일치합니다. 이러한 패턴은 CPI 예측 정확도를 높이기 위해 구성 요소 지수와 감정 점수를 모두 통합하는 것의 가치를 강조합니다.

### 3.2 뉴스 기사 수집 및 감정 추출

네이버 뉴스(<https://www.news.naver.com>)에서 'CPI'를 검색 키워드로 사용해 CPI 관련 뉴스 기사를 수집했습니다. 이 기사들은 감정 분석에서 외부 사건을 표현하는 데 사용되었습니다. 코퍼스는 2010년 1월 1일부터 2021년 11월 21일 사이에 게시된 기사로 구성되었습니다. 각 기사에는 발행일, 제목, 본문, 출처가 함께 제공됩니다. 네이버 뉴스는 다양한 국내 언론사의 기사를 한데 모아 다양한 뉴스 콘텐츠의 중앙 저장소를 제공하는 플랫폼입니다. 수집된 총 기사 수는 4,동안 213,983건입니다. 수집된 데이터는 감정 분석과 지수 계산을 통해 단일 변수( )로 가공되어 CPI 예측에 활용되었습니다.

**3.2.1 TextRank를 사용한 요약.** 감정 지수를 계산할 때 전체 기사를 분석하면 관련 없는 내용이 감정 지수에 포함되는 경우가 많아 핵심 메시지 분석이 복잡해지기 때문에 TextRank [44]를 사용하여 각 기사에 대한 요약을 작성했습니다. 미할시아 외[44]의 제안에 따라 기사당 3개의 핵심 문장을 추출했습니다.

TextRank는 추출 요약에 효과적으로 사용할 수 있는 텍스트 처리를 위한 비지도, 그래프 기반 랭크 모델입니다[45]. 이 모델은 텍스트에 있는 문장을 그래프로 표현하는 방식으로 작동하며, 각 문장은 노드가 되고 문장 사이의 에지는 유사도( )에 따라 가중치가 부여됩니다[46]. 그런 다음 이 알고리즘은 PageRank [47]와 유사한 그래프 기반 순위 지정 알고리즘을 적용하여 가장 중요한 문장을 식별합니다[44]. 문장 추출을 위한 TextRank 알고리즘은 다음 단계를 따릅니다 [46]: (1) 텍스트를 다음과 같이 분할합니다.

문장, (2) 문장을 노드로 하는 그래프를 만들고, (3) 문장 간의 유사성을 계산하여 가중치가 부여된 에지를 만들고, (4) PageRank [47] 알고리즘을 적용하여 문장의 순위를 매깁니다;  
(5) 가장 순위가 높은 문장을 요약으로 선택합니다.

가장 대표적인 문장만 추출함으로써 TextRank는 기사의 핵심 콘텐츠에 집중하여 관련성이 낮은 문장에서 노이즈를 줄임으로써 잠재적으로 감성 분석의 정확도를 향상시킬 수 있습니다[45].

**3.2.2 감성 지수 계산.** 총합으로부터 감성 지수를 계산하기 전에 먼저 가장 적합한 감성 분류기를 결정하기 위한 예비 연구를 수행했습니다. 이 연구는 두 가지 임베딩 기반 분류기 모델을 사용하여 한국어 번역 버전의 금융 구문 은행 데이터셋 [48]을 대상으로 수행되었습니다: KoBERT [40]와 KLUE-BERT [39].

파이낸스 프레이즈뱅크 데이터 세트[48]는 금융 전문가가 긍정, 중립, 부정으로 수동 라벨을 붙인 4,846개의 영어 문장으로 구성되어 있습니다. 번역된 데이터 세트의 라벨 비율은 긍정 28.22%, 중립 59.27%, 부정 12.51%였습니다.

KoBERT는 한국어를 위해 특별히 설계된 사전 학습된 BERT 모델입니다. 한국어 데이터 세트에 대해 학습되어 다국어 모델에 비해 한국어 텍스트와 관련된 작업에서 성능이 향상됩니다. 반면 KLUE-BERT[39]는 한국어의 고급 NLP 작업을 위해 설계된 KLUE 벤치마크의 일부입니다. 대규모 한국어 텍스트 말뭉치에 대해 훈련된 KLUE-BERT는 한국어의 뉘앙스를 효과적으로 포착합니다.

긍정, 중립, 부정 등 감성 분류는 F1 점수를 기준으로 평가되었습니다.

분류기 성능은 정확도와 회상률의 균형을 맞추는 F1 점수를 사용해 평가했습니다. 이러한 메트릭은 다음과 같이 정의됩니다:

$$\text{정밀도} = \frac{TP}{TP + FP}, \quad \text{리콜} = \frac{TP}{TP + FN}, \quad (2)$$

$$\text{F1 점수} = 2 \times \frac{\text{정밀도} \times \text{리콜}}{\text{정밀도} + \text{리콜}}, \quad (3)$$

여기서 TP, FP, FN은 각각, 오탐, 미탐을 나타냅니다. F1 점수는 정확도와 회수율을 동등하게 강조하기 때문에 불균형한 데이터 세트에 특히 유용합니다.

사전( ) 제한 실험에서 KLUE-BERT는 KoBERT보다 우수한 성능을 보였습니다[49]. 표 2는 평가 결과를 보여주는데, KLUE-BERT의 F1 점수는 0.852로 KoBERT의 0.434에 비해 월등히 높았습니다. 이를 바탕으로 주요 예측 모델의 입력으로 사용되는 감성 지수를 계산하기 위한 감성 분류기로 KLUE-BERT를 채택하게 되었습니다.

2010년 1월 1일부터 2021년 11월 1일까지의 208,455건의 CPI 관련 뉴스 아티클에 대해 KLUE-BERT를 사용하여 감성 분석을 수행했습니다. 이 분석은 기사를 긍정 70,, 중립 70,348건, 부정 67,783건으로 분류하여 데이터 세트의 감성 프로필을 제공했습니다.

그런 다음 데이터 세트의 각 뉴스 기사를 긍정, 중립, 부정으로 분류하여 일일 감성 지수를 구성했습니다. 각 날짜의 감성 지수는 다음을 사용하여 계산했습니다.

표 2. 감성 분류기 모델에 대한 평가 결과.

사전 학습된 모델	정확성	정밀도	리콜	F1 점수
KLUE-BERT	0.852	0.852	0.852	0.852
KoBERT	0.436	0.505	0.436	0.434

<https://doi.org/10.1371/journal.pone.0321530.t002>

Antweiler 등이 [50]에서 처음 제안한 공식으로,センチメント 분석 연구에 널리 적용되었습니다[14,51,52]:

$$\text{감정 점수} = \frac{N_t^{\text{pos}} - N_t^{\text{neg}}}{N_t^{\text{pos}} + N_t^{\text{neg}}}, \quad (4)$$

여기서  $N_t^{\text{pos}}$  및  $N_t^{\text{neg}}$ 는 각각 긍정 및 부정 뉴스 항목의 수입니다.

일  $t$ 에서 -1에서 +1 사이로 묶인 감정 점수는 일일 감정 추세를 나타냅니다.

**3.2.3 CPI 예측 지표로서의 감정 지수에 대한 평가.** 이 분석에서는 감정 지수를 계산하는 방법론이 이론적으로 타당 한지, 감정 지수가 CPI 움직임을 예측하는 신뢰할 수 있는 선행 지표로 작용할 수 있는지는 두 가지 핵심 질문을 통해 감정 지수의 CPI 움직임 예측 효과를 .

이러한 의문을 해결하기 위해 세 가지 관계에 대한 교차 상관관계 분석을 실시했습니다.

먼저, 측정 방법론을 검증하기 위해 코로나19 확진자 수와 감정 지수 간의 관계를 조사했습니다. 그런 다음 감정 지수를 후행 총 CPI와 비교하여 예측 능력을 평가하고, 후행 감정 지수에 대한 총 CPI를 분석하여 피드백 효과를 조사했습니다.

두 시계열  $x_t$ 와  $y_t$ 사이의 교차 상관 계수는 다음과 같이 정의됩니다:

$$\rho_{xy}(\tau) = \frac{\text{Cov}(x_{t-\tau}, y_t)}{\sqrt{\text{Var}(x_{t-\tau}) \cdot \text{Var}(y_t)}} = \frac{\gamma_{xy}(\tau)}{\sigma_x \sigma_y}, \quad (5)$$

여기서  $x_{t-\tau}$ 는 시계열  $x$ 의 지연된 값( $\tau$  시간 단계)을 나타내고,  $y_t$ 는 현재 값을 나타냅니다. 노이즈가 있는 원본 데이터에 5개 월 이동 평균 평활화를 적용하여 장기적인 패턴을 보다 효과적으로 포착할 수 있었습니다.

그림 3과 그림 4(a) 및 그림 4(b)에서 x축은 시간 지연을 나타내고 y축은 해당 상관관계 계수 값을 나타냅니다. 그림 3은 다음 사이의 상관관계를 보여줍니다.

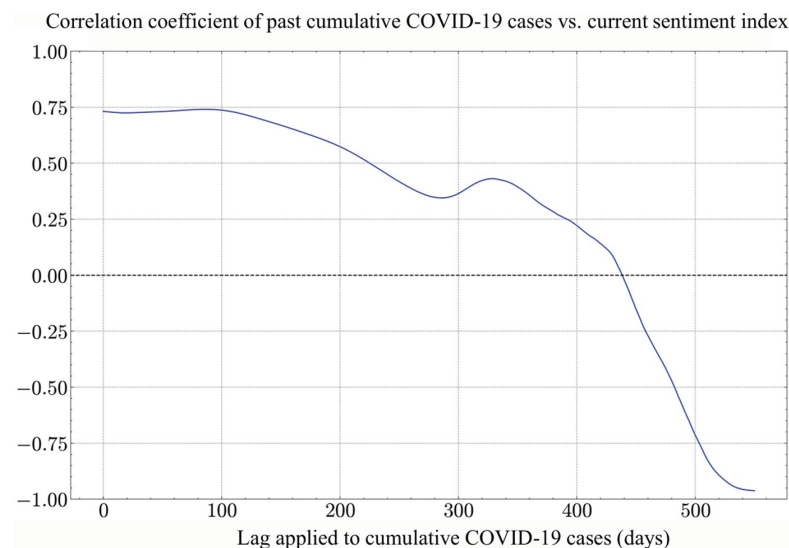
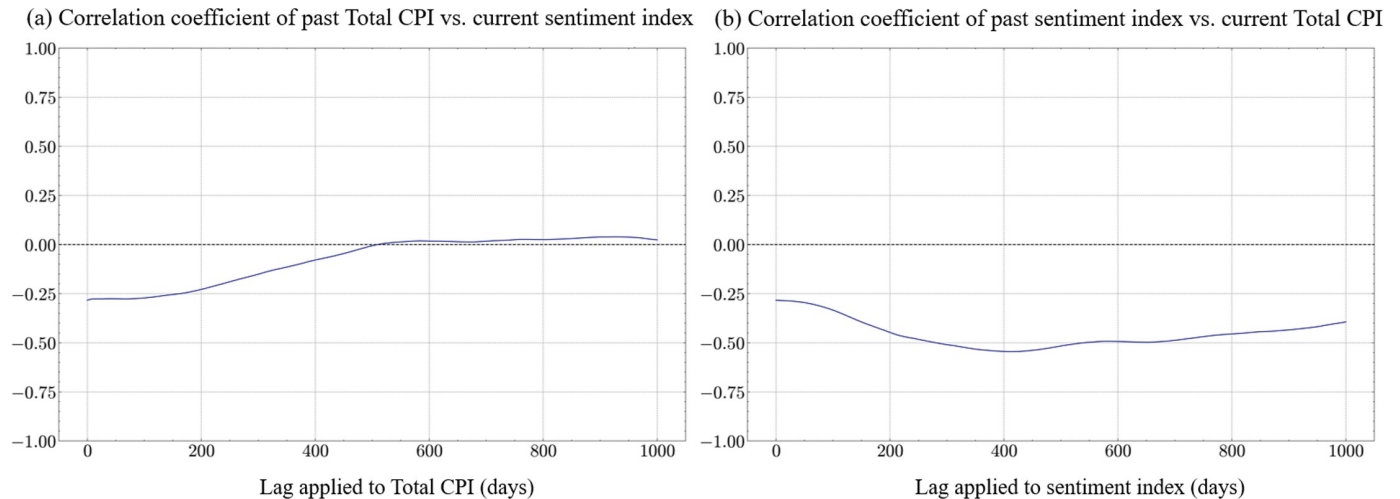


그림 3. 과거 코로나19 누적 확진자 수와 현재 감정 지수 간의 상관관계. 장기적인 부정적 추세는 정서에 대한 부정적 영향이 지연되고 있음을 시사합니다.

<https://doi.org/10.1371/journal.pone.0321530.g003>



**그림 4. 시간 경과에 따른 총 CPI와 감정 지수의 상관관계.** (a) 과거 총 CPI와 현재 감정 지수 간의 상관관계는 전반적으로 약합니다. (b) 과거 감정과 현재 총 CPI 간의 상관관계는 음의 상관관계를 유지하고 있습니다.

<https://doi.org/10.1371/journal.pone.0321530.g004>

과거 누적 COVID-19 사례와 적용된 시차에 따른 감정 지수를 일 단위로 측정된 함수입니다. 그림 4(a)와 그림 4(b)는 각각 과거 총 CPI와 현재 감정 지수 간의 상관관계, 과거 감정 지수와 현재 총 CPI 간의 상관관계를 일 단위의 시차에 따른 함수로 나타냅니다.

그림 3은 감정 지수와 지연 누적 확진자 수 간의 상관관계가 지연 기간에 따라 크게 달라지며, 최대 550일의 지연이 있음을 보여줍니다. 시차가 0일인 경우 상관관계는 계수 +0.80으로 강한 양의 관계를 나타내며, 이는 최근 코로나19 확진자 수가 많을수록 감정 지수가 높다는 것을 의미합니다. 이 긍정적인 관계는 지속되지만 시간이 지남에 따라 약화됩니다. 480일이 지나면 상관관계가 0을 넘고, 500일 이상으로 시차가 길어지면 관계가 음의 관계가 되어 500일에는 약 -0.75에 도달하고 550일에는 -0.80 미만으로 급락합니다. 단기적으로는 감정이 사례 수와 양의 상관관계가 있는 것이 직관적이지 않은 것처럼 보일 수 있지만, 이러한 패턴은 다음과 같이 설명할 수 있습니다.

대중의 인식 제고, 미디어 내러티브, 발병 당시 정서를 형성한 정부 조치 등의 요인에 영향을 받을 수 있습니다. 그러나 시간이 지날수록 과거 사례 수는 현재 정서와 반비례하는 것으로 나타나 부정적인 연관성이 지연되는 것으로 나타났습니다. 주목할 만한 점은 코로나19 누적 확진자 수가 총 CPI와 강한 상관관계(시차 0에서 +0.94)를 보이지만, 팬데믹의 일시적인 특성과 목표 데이터 세트 길이로 인해 이 관계를 주요 예측 변수로 제외했다는 점입니다.

그림 4(a)에서 감정 지수와 0~1000일의 시차를 둔 총 CPI 간의 상관 계수는 너무 약해서 유의미한 것으로 간주할 수 없습니다. 0일의 시차에서 상관관계는 약 -0.3으로, 특정 날짜에 총 CPI가 높으면 같은 날 감정이 낮아지는 경향이 있음을 나타냅니다. 시차가 200일로 증가하면 상관관계는 약 -0.15로 개선되어 음의 관계가 약해집니다. 400일이 지나면 상관관계가 0.0에 가까워져 과거 총 CPI 값이 현재 심리에 미치는 영향이 시간이 지남에 따라 감소한다는 것을 나타냅니다. 600일에서 800일 사이에는 상관관계가 -0.05에서 0.0 사이에서 약간 변동하고, 1000일이 지나면 0.0 근처에서 안정화됩니다. 이 패턴은 최근의 총 CPI 값이 심리와 음의 상관관계를 보이지만, 시차가 길어질수록 이 효과는 약해진다는 것을 시사합니다. 약 400일이 지나면 총 CPI의 과거 값은 현재 감정과 측정 가능한 관계가 거의 또는 전혀 없습니다.

그림 4(b)에서 상관관계는 모든 시간 지연에 걸쳐 일관되게 음의 값을 유지합니다. 지연이 0인 경우 상관 계수는 약 -0.28입니다. 지연이 증가함에 따라 상관관계는 더욱 음의 값을 가지며, 약 400일에서 -0.5에 가까운 최저점에 도달합니다. 이 시점 이후에는 상관관계가 약간 상승하기 시작하지만 여전히 음의 값을 유지하여 과거 심리와 현재 CPI 사이에 지속적인 역관계가 있음을 나타냅니다. 이러한 상관관계에 대한 계산된 p-값은 모두 0.01 미만으로 95% 또는 99%와 같은 일반적인 신뢰 수준에서 통계적 유의성이 있음을 확인했습니다. 이는 관찰된 상관관계가 무작위적인 변동에 의한 것일 가능성이 낮다는 것을 의미합니다. 음의 상관관계는 감성 감성어 값이 높을수록 과거에 낙관적인 심리는 미래의 낮은 CPI 값과 연관되는 경향이 있습니다. 이는 낙관적 심리가 인플레이션이 낮은 시기에 선행하는 반면, 비관적 심리는 높은 인플레이션 압력에 의해 낮아진다는 것을 나타낼 수 있습니다. 400일 시차를 두고 관찰된 가장 강력한 관계는 심리의 변화가 가격 수준에 영향을 미칠 수 있는 시간대를 시사합니다. 즉, 과거 심리지수는 미래 CPI와 유의미한 반비례 관계를 보이며, 400일 정도의 시차를 두고 가장 뚜렷한 효과가 나타납니다.

요약하면, 감정 지수를 계산하는 데 사용된 방법론이 경험적으로 타당하며, 감정 지수와 총 CPI 사이의 안정적인 음의 상관관계는 감정 지수가 CPI 움직임을 예측하는 유용한 지표가 될 수 있음을 나타냅니다. 이는 5.3.2절의 학습 중 감정 지수를 사용하거나 사용하지 않은 손실 환경을 비교함으로써 더욱 검증됩니다.

### 3.3 데이터 전처리 및 증강

CPI와 뉴스 감성 지수는 모두 일관된 배율을 보장하기 위해 정규화됩니다. 이를 위해

모든 값은 다음을 사용하여 0-1 범위로 정규화되었습니다: 
$$x = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})}$$
 감성 점수는

7일간의 평활화 과정. 시간적 무결성을 유지하고 데이터 유출을 방지하기 위해 학습 데이터 세트와 테스트 데이터 세트를 별도로 처리했습니다.

우리는 월별 관측치를 일별 값으로 변환하는 선형 보간을 통해 CPI 지수를 보강했습니다. 선형 보간은 월별 데이터에서 일별 데이터를 생성하면서 편향 없이 원본 데이터의  $e_s$  형태를 보존할 수 있습니다[53,54]. 표 3은 원본 데이터와 증강 데이터에 대한 자세한 통계를 보여줍니다. 그림 5(a)는 원본 CPI 데이터를 보여줍니다.

그림 5(b)는 선형 보간을 사용하여 증강된 일별 데이터를 보여줍니다. 비교를 통해 선형 보간된 데이터가 원본과 동일한 추세를 유지하고 있음을 알 수 있습니다.

선형 보간에서 점  $p_1$ 과  $p_2$ 의 함수 값이 각각  $f(p_1)$ 과  $f(p_2)$ 인 경우,  $p_1$ 과  $p_2$ 사이의 임의의 점  $p$ 에서의 추정값  $f(p)$ 는 다음과 같이 주어집니다:

$$F(P) = \frac{D_2}{d_1 + d_2} F(P_1) + \frac{D_1}{d_1 + d_2} F(P_2) \quad (6)$$

여기서  $d_{(1)}p$ 과  $p_1$ 사이의 거리이고,  $d_2$ 는  $p$ 와  $p_2$ 사이의 거리입니다.

표 3. 원본과 증강 콘텐츠에 대한 요약 통계.

데이터 세트	분할	Count	평균	St.	Med.	Min.	Max.	시작 날짜	종료 날짜
증강	기차	3042	98.29	3.65	99.11	89.97	104.34	2010-01-01	2018-04-30
	Val.	365	104.67	0.46	104.59	103.93	105.65	2018-05-01	2019-04-30
	테스트	915	106.15	1.34	105.56	104.56	109.48	2019-05-01	2021-10-31
원본	기차	100	98.22	3.68	99.09	89.97	104.29	2010-01-01	2018-04-01
	Val.	12	104.64	0.49	104.59	103.93	105.65	2018-05-01	2019-04-01
	테스트	30	106.08	1.29	105.55	104.56	108.95	2019-05-01	2021-10-01

<https://doi.org/10.1371/journal.pone.0321530.t003>

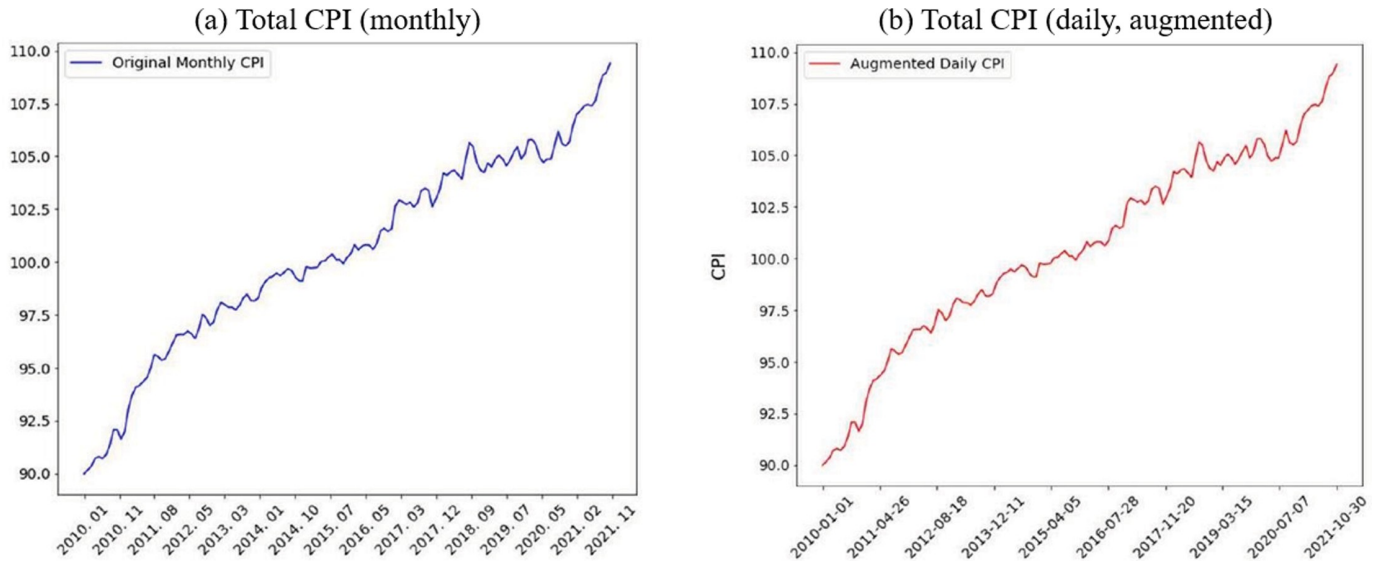


그림 5. (a) 월별 빈도의 원본 CPI 데이터, (b) 선형 보간을 통해 생성된 일별 빈도 데이터. 선형 보간은 데이터를 보강하여 추세와 패턴을 보존하면서 딥러닝 모델에 적합합니다.

<https://doi.org/10.1371/journal.pone.0321530.g005>

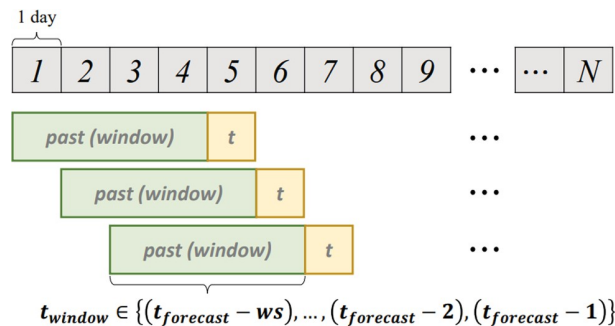


그림 6. 슬라이딩 윈도우 데이터 세트 구성 프로세스.

<https://doi.org/10.1371/journal.pone.0321530.g006>

이 보강 작업의 결과로 원래 143개월의 데이터는 4,323개의 일일 데이터 포인트로 확장되었습니다. 표 3에는 증강된 데이터 세트와 원본 데이터 세트의 요약 통계가 나와 있습니다. 간결성을 유지하기 위해 총 CPI 값만 제시했습니다.

각 월의 일별 예측값을 평균화하여 상관관계가 있는 실제 보고된 월별 값과 비교했습니다.

### 3.4 슬라이딩 창 데이터 집합 만들기

데이터 준비를 위해 Chou 등[14]의 방법론에 따라 슬라이딩 윈도우 데이터 세트를 구축했습니다. 일관성을 보장하기 위해 이들의 연구에서 사용한 것과 동일한 10개월의 룩백 윈도우를 채택했습니다.

그림 6과 같이 고정된 크기의 창이 시계열을 가로질러 단계적으로 이동하면서 과거 관측치(녹색)를 입력으로 캡처하고 이를 목표 값(노란색)과 연결합니다. 입력 시퀀스는  $t_{\text{예측}} - ws$ 에서  $t_{\text{예측}} - 1$ 에 걸쳐, 여기서  $ws$ 는 창 크기입니다. 이 접근 방식은 시간적 종속성을 유지하여 효과적인 예측을 용이하게 합니다.



일단 구축된 데이터 세트는 학습, 검증, 평가 세트로 분할됩니다. 전체 데이터 세트는 2010년 1월 1일부터 2021년 11월 1일까지 143개월에 걸쳐 있습니다. 데이터의 75%를 차지하는 훈련 기간은 2010년 1월 1일부터 2018년 4월 30일까지입니다. 검증 기간은 데이터의 10%를 차지하며 2018년 5월 1일부터 2019년 4월 30일까지입니다. 나머지 15%의 데이터를 대상으로 하는 테스트 기간은 2019년 5월 1일부터 2021년 10월 31일까지입니다. 이러한 시간적 분리를 통해 모델은 과거 데이터로 학습되고, 중간 기간에 검증되며, 가장 최근 데이터로 테스트되어 견고한 평가를 받을 수 있습니다.

### 3.5 예측을 위한 CNN-LSTM 모델 제안

CPI 관련 수치 특징과 감정 점수는 1차원(1D)이므로, 1D CNN 레이어를 사용하여 특징 추출을 수행했습니다. 하이퍼파라미터는 [30]에 따라 커널 크기 2, 보폭 1로 구성했습니다. 커널 크기가 2인 경우, 1D CNN 레이어에는 두 개의 가중치 파라미터( $w_1, w_2$ )와 하나의 바이어스 파라미터( $b$ )가 포함됩니다. 그림 7은 커널이 각 반복에서 단계 크기 1로 왼쪽에서 오른쪽으로 이동하는 커널 시프트 프로세스를 보여 주며, 입력 값  $x_t$  및  $x_{t+1}$ 의 특징이 데이터에서 어떻게 추출되는지 보여줍니다. 중간 특징 점  $z_1$ 과  $z_{(2)}$ 를 계산하는 공식은 다음과 같습니다.

$$Z_1 = W_1 \cdot X_t + W_2 \cdot X_{t+1} + B \quad (7)$$

$$Z_2 = W_1 \cdot X_{t+1} + W_2 \cdot X_{t+2} + B \quad (8)$$

여기서  $d_{(1)} \in x_t$  및  $x_{t+1}$ 에서 추출한 특징이고,  $d_{(2)} \in x_{t+1}$  및  $x_{t+2}$ 에서 추출한 특징입니다. 가중치  $w_1$  및  $w_2$ 는 커널 내의 입력에 적용되며,  $b$ 는 가중치 합에 추가된 바이어스 항입니다. 5.3.1절에서 이 계층이 피쳐에 미치는 영향에 대해 설명합니다.

1D CNN 레이어에서 특징을 추출한 후 LSTM 레이어는 정보를 순차적으로 처리합니다. 먼저 잊음 게이트 계층에서 특정 정보를 버려야 하는지 여부를 평가합니다. 다음으로, 입력 게이트 레이어를 통해 불필요한 정보를 제거하고

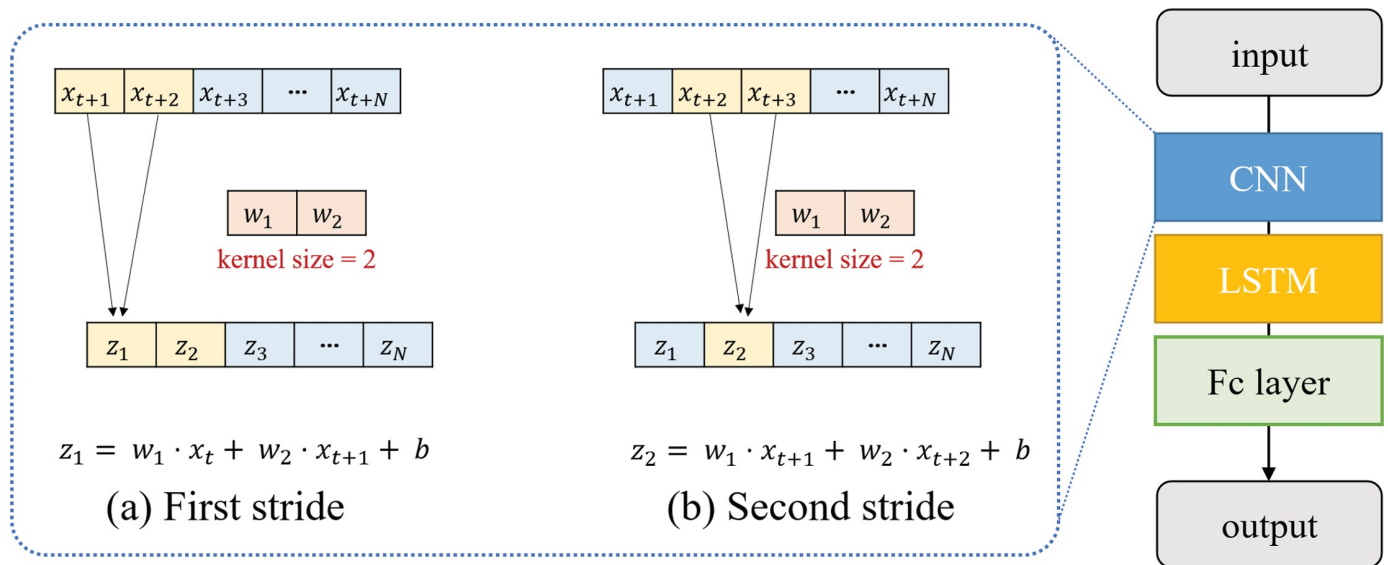


그림 7. 시계열 예측을 위해 제안된 CNN-LSTM 모델. 왼쪽 패널은 커널 크기 2의 1D CNN 스트레이드를 통한 특징 추출을 보여줍니다. 오른쪽 패널은 로컬 패턴을 위한 CNN, 장기 종속성을 위한 LSTM, 예측을 위한 완전히 연결된 레이어를 결합한 모델 아키텍처를 보여줍니다.

<https://doi.org/10.1371/journal.pone.0321530.g007>



셀 상태가 그에 따라 업데이트됩니다. 마지막으로 출력 게이트 레이어는 업데이트된 셀 상태 정보를 기반으로 실제 예측 결과를 생성합니다.

$$\text{포켓 게이트: } f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (9)$$

$$\text{입력 게이트: } i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (10)$$

$$\text{입력 변조 게이트: } g_t = \tanh(W_g x_t + U_g h_{t-1} + b_g) \quad (11)$$

$$\text{업데이트된 셀 상태: } c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (12)$$

$$\text{출력 게이트: } o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (13)$$

$$\text{메모리 셀의 출력: } h_t = o_t \odot \tanh(c_t), \quad (14)$$

여기서  $\sigma$ 는 원소별 시그모이드 함수를 나타냅니다.  $\tanh$ 은 원소별 쌍곡선 접선 함수를 나타냅니다.  $\odot$ 은 하다마드(원소별) 곱을 나타냅니다.

$W_f, W_g, W_i, W_o$ 는 입력 가중치 행렬이고,  $U_f, U_g, U_i, U_o$ 는 반복 가중치 행렬입니다.

$b_f, b_g, b_i, b_o$ 는 바이어스 벡터입니다.

## 4 실험 설정

### 4.1 입력 및 예측 지평선 설정

시계열 예측의 경우, 적절한 시간 창 설정을 결정하는 것이 모델 성능에 매우 중요합니다. 저희 모델은 10개월의 룩백 윈도우를 사용하여 각 예측에 대해 310일(31일 × 10개월)의 과거 데이터를 처리했습니다. 이 기간의 길이는 다양한 과거 기간이 예측 정확도에 미치는 영향을 평가하는 경험적 실험을 통해 결정되었습니다. 기간이 짧으면 과거 패턴을 충분히 포착하지 못해 정보 손실이 발생하고, 기간을 10개월 이상으로 늘리면 예측 정확도가 의미 있게 개선되지 않고 계산 복잡성만 증가했습니다.

이 모델은 1개월(31일) 동안의 예측을 생성하고, 일일 예측을 평균화하여 공식 CPI 발표 주기와 일치하는 월별 값을 생성했습니다. 이러한 평균화 접근 방식은 모델의 일별 세분성과 CPI 발표의 월별 주기 사이의 불일치를 해결하여 의미 있는 성과 평가를 보장하는 동시에 일일 시장 역학을 포착하는 모델의 기능을 유지했습니다.

### 4.2 하이퍼파라미터 설정

Optuna 기반 베이지안 최적화[55]와 그리드 검색의 조합을 사용하여 체계적인 하이퍼파라미터 튜닝을 수행했습니다.

Optuna는 더 빠르고 자동화된 최적화를 가능케 하지만, 본질적으로 해석 가능성과 검색 프로세스에 대한 직접적인 제어를 감소시킵니다. 이 문제를 해결하기 위해 하이브리드 전략을 채택했습니다: 해석 가능성을 유지하기 위해 그리드 검색을 통해 조정된 숨겨진 단위 크기를 제외한 모든 파라미터에 Optuna를 사용했습니다.

훈련의 확률적 특성을 고려하여 각 구성은 20개의 런덤 시드를 사용하여 평가했습니다. 편차를 줄이기 위해 중간 순위의 결과 5개를 기준으로 결정을 내렸습니다. 탈락률은 0.1~0.7 범위 내에서 최적화되었으며, Optuna는 0.6으로 수렴하는 최적 설정으로 정규화와 성능의 균형을 맞췄습니다. 빠른 수렴과 안정성 사이의 균형을 맞추기 위해 학습 속도도 Optuna를 통해 최적화되었습니다.

CNN 하이퍼파라미터의 경우 [30]을 따라 커널 크기를 2로, 보폭을 1로 설정했습니다. 숨겨진 단위 크기는 64, 128, 256, 512 단위의 그리드 검색을 사용하여 개별적으로 조정한 결과 512 단위가 가장 우수한 성능을 제공했습니다. 단위 수가 적은 모델은 장기 종속성을 포착하는 데 어려움을 겪었고, 512개 이상의 대규모 아키텍처는 의미 있는 개선 없이 계산 비용이 증가했습니다.

훈련 최적화는 적응형 학습 속도와 안정성을 위해 선택된 아담 옵티마이저 [56]를 사용하여 수행했습니다. 과적합을 방지하기 위해 인내심을 1로 설정하여 검증 성능이 정체되면 학습을 중단하는 조기 중지를 구현했습니다. 이러한 보수적인 접근 방식은 불필요한 계산을 피하면서도 효과적인 학습을 가능하게 했습니다. 모든 기준 모델은 공정한 비교를 위해 이러한 최적화된 하이퍼파라미터를 사용하여 구성되었습니다.

### 4.3 평가 지표

몇 가지 지표를 사용해 모델의 성능을 평가합니다: RMSE, 평균 절대 오차(MAE), 평균 절대 백분율 오차(MAPE), 대형 평균 절대 백분율 오차(SMAPE)가 그것입니다. 다음 공식에서  $n$ 은 테스트 데이터 포인트의 총 개수,  $y_{(i)}$ 는 테스트 데이터의 실제 값,  $\hat{y}_{(i)}$ 는 테스트 데이터의 예측값을 나타냅니다. RMSE는 제곱 잔차의 평균의 제곱근을 계산합니다. 공식은 다음과 같습니다:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (15)$$

MAE는 RMSE보다 이상값에 덜 민감하기 때문에 더 견고할 수 있습니다. 공식은 다음과 같습니다:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (16)$$

MAPE는 실제 값과 예측 값의 차이를 실제 값으로 나누고 실제 값에 대한 오차의 상대적 비율을 0에서 100% 사이로 계산하여 계산합니다. 그런 다음 이전에 계산된 값의 절대값을 평균합니다. 공식은 다음과 같습니다:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (17)$$

실제 값이 0인 경우 MAPE를 계산할 수 없는 문제를 해결하기 위해 개발된 것이 SMAPE입니다. 또한 SMAPE는 MAPE의 절대값이 동일한 오차를 가지더라도 실제 값과 예측 값의 관계를 고려하여 과대평가를 제한합니다. 공식은 다음과 같습니다:

$$\text{SMAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|) / 2} \times 100 \quad (18)$$

정규화된 루트 평균 제곱 오차(NRMSE)는 스케일과 무관한 측정값을 제공합니다.

평가 데이터 세트의 표준 편차에 대해 RMSE를 정규화하여 오류를 확인합니다. 이 정규화를 통해 척도가 다른 데이터 세트 간에 직접 비교할 수 있습니다. NRMSE의 공식은 다음과 같습니다:

$$\text{NRMSE} = \frac{\text{RMSE}}{\text{stddev}_y} \quad (19)$$

여기서  $\text{stddev}_y$ 는 평가 데이터 세트의 실제 값의 표준 편차입니다. 목표 변수(총 CPI)의 표준 편차로 RMSE를 확장함으로써 NRMSE는 예측 오류에 대한 표준화된 평가를 제공합니다. 이 조정은 CPI의 자연 변동성과 관련된 모델 성능을 맥락화하여 오차의 실질적인 중요성에 대한 보다 명확한 관점을 제공합니다.

표 4. 평가 범위 통계 요약.

통계	가치
평균	105.981
중앙값	105.512
표준 편차	1.212

<https://doi.org/10.1371/journal.pone.0321530.t004>

이 연구에서 CPI 평가 세트(룩백 기간을 제외한 테스트 세트)의 표준편차는 표 4와 같이 1.212로, 데이터의 내재적 변동을 나타냅니다. 이 임계값에 가깝거나 그보다 낮은 NRMSE 값을 가진 모델은 실제 CPI 변동을 효과적으로 포착하여 실질적인 예측 유용성을 입증합니다.

4.4 전산 설정

실험 프레임워크는 Python 3.9로 개발되었으며, 딥 러닝 구현을 위해 PyTorch 라이브러리를 활용했습니다. 계산 인프라는 AMD EPYC 9654 96코어 프로세서 및 1TB RAM과 결합된 NVIDIA L40S GPU로 구성되었습니다. 모델을 비교하는 동안 최고 수준의 수치 정밀도를 유지하기 위해 모든 계산은 고정밀도 부동 소수점(float64) 연산을 사용하여 수행되었습니다.

5 실험 결과

5.1 기준 방법과의 성능 비교

5.1.1 베이스라인 모델에 대한 설명. 제안된 모델의 효과를 평가하기 위해 표준 머신러닝 및 딥러닝 기준선과 성능을 비교합니다. 이 평가에는 모델의 변형과 기준 예측 접근 방식이 모두 포함됩니다.

세 가지 버전의 축소된 모델은 다양한 구성 요소의 기여도를 분석하기 위한 내부 기준선으로 사용됩니다. *wo. CPI-multi* 변형은 총 CPI만 입력으로 사용하여 모든 보조 CPI 관련 변수를 제외합니다. *sent*. 변형은 감정 정보를 제거하여 예측 정확도에 미치는 영향을 평가합니다. *wo. aug.* 변형은 데이터 보강을 생략하고 CPI 관련 변수와 감정 점수를 모두 유지하여 모델 성능에 대한 보강의 영향을 분리할 수 있습니다.

외부 기준선의 경우, 다변량 시계열 예측에 일반적으로 적용되는 네 가지 머신 러닝 모델을 통합합니다. 올가미 회귀[43]는 *L1* 정규화를 통한 특징 선택 기능으로 가장 영향력 있는 CPI 구성 요소를 식별하는 데 도움이 됩니다. 그러나 선형 관계()를 가정하기 때문에 복잡한 종속성을 포착하는 데는 한계가 있습니다. 지원 벡터 회귀(SVR)[57]는 커널 함수를 통해 비선형 관계를 모델링할 수 있기 때문에 선택되지만, 계산 비효율성과 시계열 모델링의 부족으로 대규모 시계열 예측에는 적합하지 않습니다. 랜덤 포레스트[58]는 과적합에 대한 견고성과 다양한 변수 간 상호작용을 포착할 수 있는 능력 때문에 선택되었습니다. 랜덤 포레스트는 구조화된 특징을 효과적으로 처리하지만, 지속적으로 성장하는 의사 결정 트리()에 의존하기 때문에 시간적 종속성을 모델링하는 데는 한계가 있습니다.

구조화된 회귀 작업에서 강력한 성능을 발휘하는 최적화된 그래디언트 부스팅 알고리즘인 XGBoost[59]가 포함되어 있으며, 반복적인 잔차 최소화를 활용하여 정확도를 향상시킵니다. 그러나 XGBoost는 본질적으로 순차적 종속성을 모델링하지 않기 때문에 원시 시계열 데이터에는 유연성이 떨어집니다.

딥러닝 기준선은 다양한 아키텍처가 CPI 예측에 미치는 영향을 평가하고 해당 구성 요소의 영향을 분석하기 위해 선택됩니다. 피드포워드 신경망(FNN)은 비선형 모델링의 기준선을 제공하지만 메모리 메커니즘이 없기 때문에 시퀀스 정보를 보유할 수 없습니다( ). 게이트 순환 단위(GRU) [60]은 LSTM[25]에 비해 계산적으로 더 효율적인 대안을 제공하여 효율성과 장기 의존성 모델링 간의 비교를 가능하게 합니다. LSTM 네트워크는 장기간에 걸쳐 의존성을 유지할 수 있다는 점에서 시계열 예측에 널리 사용되는 벤치마크로 포함됩니다. 또한, CNN 기반의 코널 특징 추출을 주의 메커니즘으로 대체하는 AttentionLSTM[61]을 평가합니다. 이를 통해 직접 다변량 주의가 CPI 데이터의 종속성을 포착하기 위한 컨볼루션 특징 추출( )의 대안이 될 수 있는지 평가할 수 있습니다.

**5.1.2 성능 분석.** 표 5는 5.1.1절에 설명된 대로 올가미, 랜덤 포레스트, XGBoost, SVR 및 피드포워드 신경망(FNN), GRU, LSTM, AttentionLSTM 등의 딥러닝 모델 등 다양한 기준 방법과 저회 모델을 비교한 것입니다.

우리 모델(전체)의 성능은 머신러닝(ML) 및 딥러닝(DL) 모델에 비해 RMSE, NRMSE, MAE, MAPE 및 SMAPE 측면에서 상당한 개선이 이루어졌음을 보여줍니다. 특히 RMSE는 ML 모델에 비해 96.27%, DL 모델에 비해 66.19% 감소했습니다. ML 모델과 비교했을 때, Ours(전체)는 97.27% 감소했습니다.

MAE에서 97.26%, MAPE에서 97.26%, SMAPE에서 97.34% 감소했습니다. DL 모델과 비교했을 때도 여전히 MAE는 71.70%, MAPE는 71.53%, 71.67% 감소하는 등 강력한 우위를 유지합니다. 그림 8은 풀 구성으로 제안한 CNN-LSTM 모델의 예측값이 다른 모델보다 실제 값에 더 가깝다는 것을 보여줍니다.

절대적인 측면에서 보면, Ours(전체)는 ML 모델에 비해 MAE 5.198, MAPE 4.875, SMAPE 5.016의 감소를 달성합니다. DL 모델과 비교했을 때, 감소율은 다음과 같습니다.

는 MAE에서 0.370, MAPE에서 0.345, SMAPE에서 0.347로 감소하여 여전히 유의미하지만 더 작습니다. 또한 Ours(전체)는 ML 모델에 비해 NRMSE가 4.279 감소하여 상당히 개선된 것으로 나타났습니다. DL 모델과 비교했을 때도 0.325 감소로 여전히 상당한 개선이 이루어졌습니다.

이러한 결과는 특히 평균 96% 이상의 감소율을 보인 ML 모델과 비교했을 때 당사의 모델(전체)의 우수성을 강조합니다. DL 모델과 비교해도 이 모델은 RMSE가 66%나 개선되었습니다.

표 6에서 볼 수 있듯이 월콕슨 사인-랭크 테스트 [62] 결과, 당사 모델과 비교 모델 간에 RMSE 성능에 차이가 없다고 가정하는 귀무가설은 0.05 미만의 유의 수준에서 기각되었습니다.

## 5.2 토론

**5.2.1 절제 연구: 주요 구성 요소의 기여도.** 표 5에 제시된 것처럼 제거 연구를 통해 개별 구성 요소(CPI-월E, 감정 지수, 데이터 증강)의 영향을 분석했습니다. 그림 9는 우리가 제안하는 완전 구성된 CNN-LSTM 모델이 제거된 모델에 비해 실제 값에 더 근접한 예측값을 생성한다는 것을 보여줍니다.

머신 러닝(ML) 모델과 딥 러닝(DL) 모델에 따라 RMSE에 미치는 다양한 구성 요소의 영향은 달라집니다. 올가미, 랜덤 포레스트, SVR 및 XGBoost에서 CPI-월E/변수를 제거해도 RMSE는 크게 변하지 않고 거의 그대로 유지됩니다. 놀랍게도 감정 데이터를 제외하면 RMSE가 약간 감소하는 소폭의 개선이 이루어집니다. 마찬가지로, 증강을 제거하면 RMSE가 약간 감소합니다.

표 5. 주요 요소(CPI-멀티, 감정 지수, 데이터 증강)를 제외했을 때의 영향을 보여주는 기준 모델과 제거 변형의 성능 비교.

	RMSE	NRMSE	MAE	MAPE	SMAPE
올가미 (wo. CPI-멀티)	7.3934	6.1002	7.3605	6.9046	7.1534
올가미 (오. senti)	6.8901	5.6848	6.8513	6.4261	6.6415
올가미(8월 둘째 주)	6.8954	5.6892	6.8570	6.4315	6.6472
올가미(전체)	7.3934	6.1002	7.3605	6.9046	7.1534
랜덤 포레스트(wo. CPI-멀티)	3.9185	3.2330	3.8772	3.6354	3.7040
무작위 숲 (오. senti)	3.9128	3.2283	3.8704	3.6290	3.6974
랜덤 포레스트(8월 둘째 주)	3.9362	3.2476	3.8887	3.6460	3.7152
랜덤 포레스트(전체)	3.9071	3.2236	3.8659	3.6249	3.6931
SVR(wo. CPI-멀티)	6.2694	5.1728	6.2380	5.8514	6.0293
SVR(wo. senti.)	6.2137	5.1268	6.1801	5.7970	5.9716
SVR (wo. aug.)	6.4050	5.2847	6.3729	5.9780	6.1637
SVR(전체)	6.2694	5.1728	6.2380	5.8514	6.0293
XGBoost(wo. CPI-멀티)	3.9783	3.2824	3.9107	3.6663	3.7370
XGBoost (wo. senti.)	3.8363	3.1652	3.7922	3.5563	3.6220
XGBoost(8월 둘째 주)	3.6482	3.0100	3.5801	3.3555	3.4148
XGBoost(전체)	3.9783	3.2824	3.9107	3.6663	3.7370
FFNN(wo. CPI-멀티)	0.3298	0.2722	0.2792	0.2613	0.2618
FFNN (wo. senti.)	0.4357	0.3595	0.3779	0.3537	0.3546
FFNN (wo. aug.)	6.0736	5.0112	6.0033	5.6283	5.8127
FFNN(전체)	0.4434	0.3659	0.3869	0.3622	0.3632
GRU(wo. CPI-멀티)	1.3995	1.1547	1.3685	1.2822	1.2908
GRU (wo. senti.)	0.3468	0.2862	0.3042	0.2865	0.2860
GRU (wo. aug.)	14.2467	11.7546	14.1988	13.3210	14.2830
GRU(전체)	0.8406	0.6935	0.7657	0.7156	0.7189
LSTM(wo. CPI-멀티)	1.6410	1.3539	1.5773	1.4766	1.4885
LSTM (wo. senti.)	0.5244	0.4327	0.4822	0.4536	0.4521
LSTM (wo. aug.)	15.6921	12.9472	15.6459	14.6795	15.8490
LSTM(전체)	0.5544	0.4574	0.4728	0.4416	0.4431
AttentionLSTM (wo. CPI-멀티)	1.5721	1.2971	1.3360	1.2460	1.2567
AttentionLSTM (wo. senti.)	0.7133	0.5886	0.5768	0.5436	0.5425
주의LSTM(8월 둘째 주)	14.4642	11.9341	14.4149	13.5236	14.5111
AttentionLSTM (전체)	0.9357	0.7721	0.8083	0.7549	0.7590
우리(wo. CPI-멀티)	0.2074	0.1711	0.1917	0.1795	0.1796
우리(오. senti)	0.2065	0.1704	0.1850	0.1743	0.1741
우리(8월)	1.4570	1.2022	1.2375	1.1540	1.1634
우리(전체)	0.2012	0.1660	0.1460	0.1373	0.1372

<https://doi.org/10.1371/journal.pone.0321530.t005>

의 약 3.1%입니다. 이러한 관찰은 ML 모델이 CPI-멀티에 크게 민감하지 않음을 나타냅니다.

하지만 감정 표현과 증강을 제거하면 약간의 개선이 나타납니다.

반대로 FFNN, GRU, LSTM, AttentionLSTM 및 제안된 모델을 포함한 DL 모델은 특정 구성 요소에 대한 의존성이 더 강한 것으로 나타났습니다. CPI-multi를 제거하면 RMSE가 73%로 크게 증가하여 그 중요성을 알 수 있습니다. 그럼에도 불구하고 감정 데이터를 제거하면 평균적으로 RMSE가 약간 개선되지만, 일부 모델의 경우 감정이 노이즈를 유발할 수 있음을 시사하는 제안된 CNN-LSTM 모델과 랜덤 포레스트는 예외입니다. 가장 두드러진 효과는 증강 제거에서 관찰되며, 이는 평균 17.45배의 RMSE 증가로 이어져 딥러닝 모델에서 강력한 성능을 보장하는 데 중요한 역할을 한다는 점을 강조합니다.

이러한 결과는 ML 모델이 CPI-멀티의 영향을 거의 받지 않으며 감정 및 증강을 제거하면 약간의 이점을 얻을 수 있음을 시사합니다. 반면, DL 모델은 증강과 CPI 멀티에 크게 의존하며, 이를 제거하면 성능이 크게 저하됩니다.

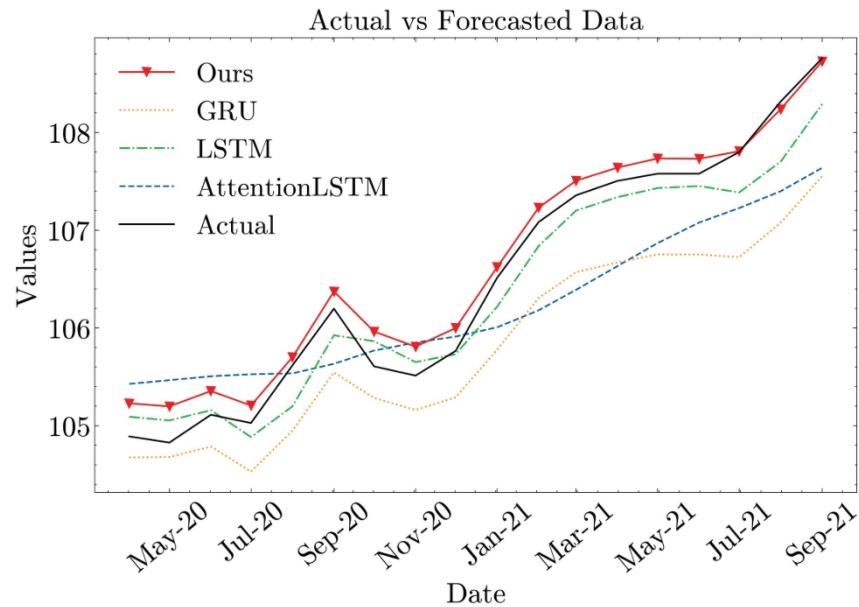


그림 8. 기본 딥러닝 방법과 모델 성능 비교.

<https://doi.org/10.1371/journal.pone.0321530.g008>

표 6. 월콕슨 서명 순위 테스트 결과.

모델	p-값
AttentionLSTM	0.037
GRU	0.037
LSTM	0.037
FFNN	0.037

<https://doi.org/10.1371/journal.pone.0321530.t006>

다른 DL 모델과 달리 감정으로 인한 모델의 성능이 향상된 것은 5.3.1절에 설명된 것처럼 제안된 CNN-LSTM 아키텍처에서 1D CNN을 사용하기 수 있으며, 이는 노이즈가 많은 감정 데이터를 부드럽게 처리하는 데 도움이 될 수 있습니다. 입력 피쳐의 노이즈를 줄임으로써 모델은 더 관련성 높은 패턴을 추출할 수 있으므로 원시 감정 데이터를 직접 처리하는 모델보다 성능이 더 뛰어납니다. 이는 또한 노이즈 완화( )가 모델의 효율성을 향상시키기 때문에 감정 데이터를 제거해도 기존 모델에 미치는 영향은 미미하지만 모델에 영향을 미치는 이유를 설명할 수 있습니다.

**5.2.2 NRMSE: 절대값과 예측 신뢰도로 RMSE를 평가합니다.** 방정식 19에 정의된 대로 NRMSE는 정규화된 RMSE로, RMSE를 평가 집합의 표준 편차로 나누어 구합니다. NRMSE가 1 미만이면 예측 오류가 있음을 나타냅니다.

가 목표 값의 변동성보다 작으면 예측이 실질적인 관련성을 유지할 수 있습니다.

이는 데이터 규모와 분산 차이를 설명하기 위해 오류 메트릭을 정규화하는 것의 중요성을 강조하는 Chai 등( [63])이 제시한 방법론과 같은 확립된 방법론과 일치합니다. NRMSE를 통합함으로써 당사의 평가는 예측 정확도에 대한 표준화된 척도를 제공하여 모델 예측과 CPI의 자연 변동 간의 직접적인 비교를 가능하게 합니다.

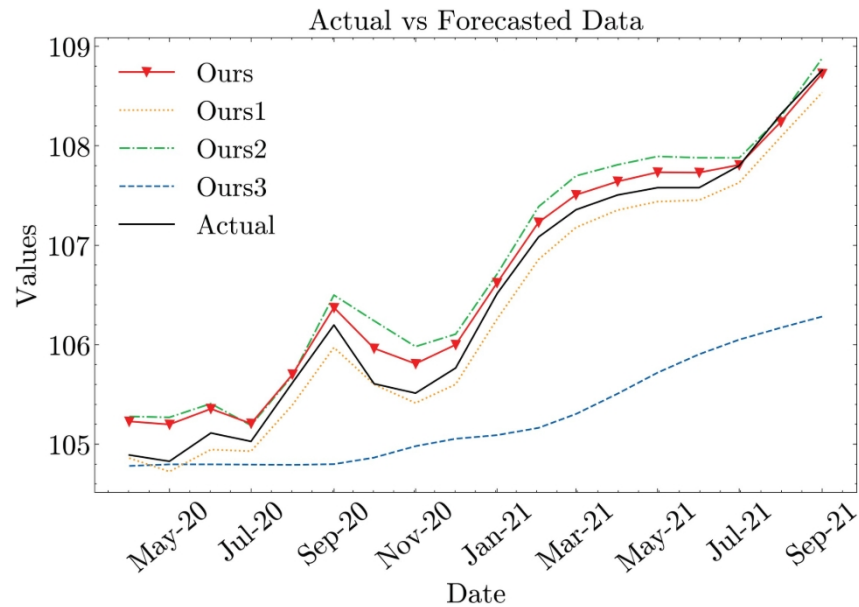


그림 9. 개별 구성 요소의 영향을 강조하기 위해 제안된 모델과 제거 변형 모델의 성능 비교. 우리의1, 우리의2, 우리의3은 각각 *CPI-멀티*, *감정 지수*, *데이터 증강이 없는* 제안 모델을 나타냅니다.

<https://doi.org/10.1371/journal.pone.0321530.g009>

이러한 접근 방식의 효과는 NRMSE 결과에 반영되어 있습니다. CNN-LSTM 모델은 평가 세트의 표준 편차의 약 6분의 1 수준인 0.1660의 NRMSE를 달성하여 CPI의 고유한 변동성에 비해 매우 정확한 예측을 생성할 수 있는 능력을 보여주었습니다. 반면, Random Forest, XGBoost, Lasso를 포함한 대부분의 머신러닝 모델은 1 이상의 NRMSE 값을 기록하여 동일한 수준의 정밀도로 CPI 추세를 포착하는 데 한계가 있음을 보여주었습니다.

궁극적으로, 우리 모델의 낮은 NRMSE는 통계적 및 실제 예측 요구 사항을 모두 해결하면서 정확하고 신뢰할 수 있는 CPI 예측을 생성하는 데 있어 실용적인 유용성을 강조합니다.

## 5.3 보충 연구

**5.3.1 1D CNN 레이어의 노이즈 감소 기능 평가: 내부 분석.** 그림 10의 중간 출력 시각화는 CNN-LSTM 모델의 내부 처리에 대한 통찰력을 제공하고 하이브리드 접근 방식의 효과를 검증합니다. CNN 계층과 LSTM 계층의 기여도를 알아보기 위해 Kim 등[64]의 방법을 적용했습니다. 그 결과 CNN 계층이 시계열 데이터의 노이즈를 효과적으로 줄이면서 필수적인 추세를 보존하는 것으로 나타났습니다.

그림 10은 감정 지수 데이터가 CNN 계층을 통과한 후 노이즈가 감소하면서 전체적인 시간적 패턴을 유지한다는 것을 보여줍니다. 감정 지수 데이터에는 불규칙한 시간적 변화가 포함되어 있기 때문에 이러한 특성을 유지하는 것은 정확한 사전 디셔너리를 위해 매우 중요합니다. 이렇게 처리된 데이터는 LSTM 계층의 입력으로 사용되어 보다 효과적인 학습을 가능하게 합니다.

**5.3.2 감정 지수의 유용성: CNN-LSTM 하이브리드 모델의 손실 환경 분석.** 손실 경관 분석은 감정 지수를 통합하면 CNN-LSTM 하이브리드 모델이 향상된다는 강력한 증거를 제공합니다. 이러한 추가는 최적화 안정성을 향상시킬 뿐만 아니라 감정 분석 작업에서 모델 성능을 강화합니다.

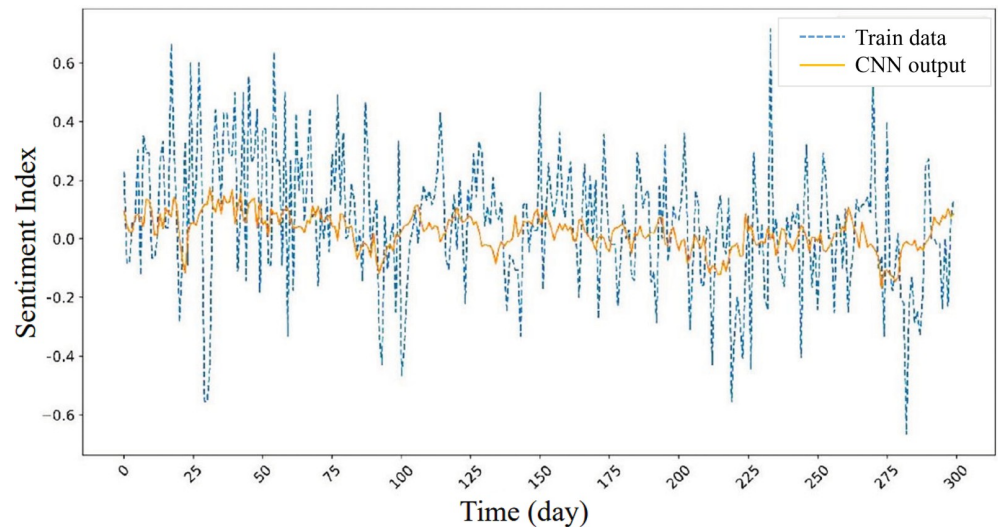


그림 10. 1D CNN 레이어의 내부 분석: 레이어를 통한 처리 전후의 입력 특징 비교.

<https://doi.org/10.1371/journal.pone.0321530.g010>

그 영향을 평가하기 위해 Li 등이 제안한 필터 정규화 방법[51]을 사용해 손실 지형을 시각화했습니다. 이 기법은 컨볼루션 필터와 완전히 연결된 뉴런을 개별적으로 노멀라이즈하여 손실 곡물을 공정하게 재현함으로써 의미 있는 비교를 용이하게 합니다.

두 가지 시나리오, 즉 감정 지수가 없는 CNN-LSTM 모델(그림 11a)과 감정 지수가 있는 모델(그림 11b)을 비교했습니다. 결과는 주요 차이점을 강조합니다. 첫째, 그림 11b의 손실 풍경이 눈에 띄게 부드러워져 보다 안정적인 최적화가 이루어졌음을 나타냅니다. 둘째, 이러한 부드러움은 경사 하강이 더 효율적으로 수렴하여 훈련을 더 강력하게 만들 수 있음을 시사합니다. 셋째, 손실 환경이 덜 혼란스러워지면 최적이지 않은 국소 최소값에 갇힐 위험이 줄어드는 동시에 잘 일반화된 솔루션에 도달 가능성이 높아집니다.

Li 등[37]은 손실 경관이 더 매끄러울수록 일반화가 더 잘되고 과적합이 줄어든다고 주장합니다. 이번 연구 결과는 이러한 관점에 부합하며, 감정 지수가 최적화 역학을 개선하고 예측 성능을 향상시키며 보다 안정적인 훈련 프로세스를 촉진한다는 것을 보여줍니다.

## 6 향후 시사점 및 작업

우리가 아는 한, 선형 보간과 같은 간단한 비모수적 방법을 사용한 시계열 증강에 대한 탐색은 제한적이었습니다. 이번 연구 결과는 이러한 간단한 접근 방식이 예측 성능을 크게 향상시킬 수 있음을 보여주며, 시계열 분석에서 경량 증강 전략의 잠재력을 강조합니다.

저희 모델은 강력한 예측 성능을 보여주지만, 몇 가지 측면에서는 더 깊은 조사가 필요합니다. 예를 들어, 관찰된 상당한 성능 향상을 고려할 때, 증강 방법에 대한 추가 연구가 예측 기술을 발전시키는 데 유용할 수 있습니다. 예를 들어, 도메인별 사전 예측을 보다 정교한 증강 전략을 통합하면 더 많은 개선 효과를 얻을 수 있습니다. 또한 거시경제 지표나 도메인 지식과 같은 보완적인 데이터 소스를 포함하면 모델의 견고성을 더욱 강화하고 데이터 기반 예측을 실제 의사 결정에 더 잘 맞출 수 있습니다.



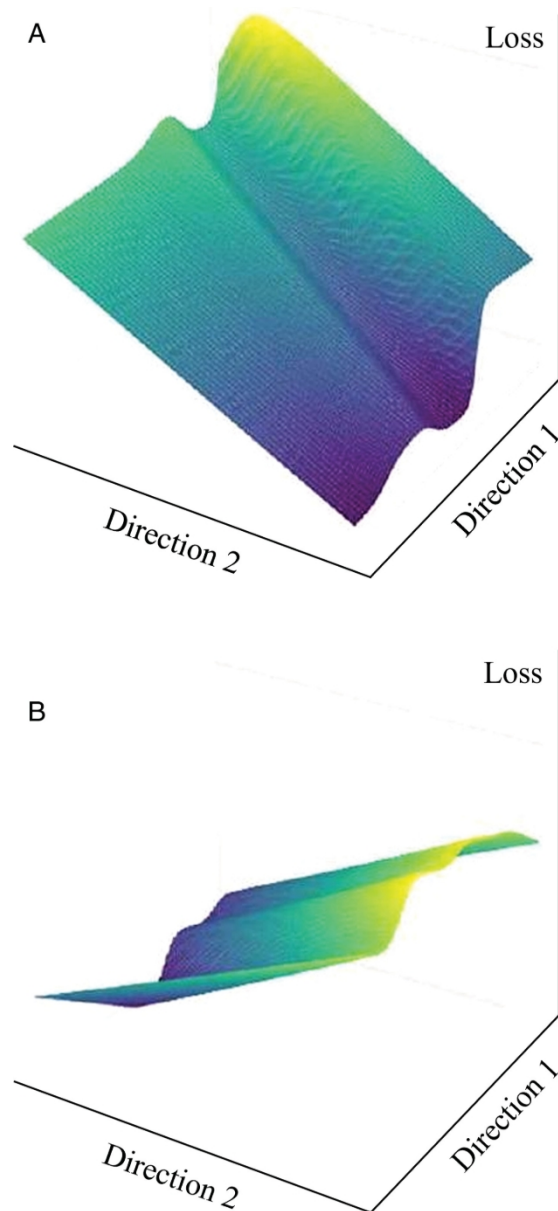


그림 11. 감정 지수가 없는 손실 표면. (a) 감정 지수가 없는 손실 표면. (b) 감정 지수가 있는 손실 표면.

<https://doi.org/10.1371/journal.pone.0321530.g011>

또한, 구성 요소 기여도의 변동성을 보다 엄격하게 조사하면 다양한 모델과 데이터 세트에서 그 영향을 더욱 명확하게 파악할 수 있습니다. 예를 들어, 향후 연구에서는 더 넓은 범위의 하이퍼 파라미터 구성과 여러 데이터 세트에서 이러한 효과의 안정성을 탐구해야 합니다. SHAP 값이나 순열 중요도와 같은 보다 체계적인 특징 속성 기법을 사용하면 예측 성능에서 각 구성 요소의 역할에 대한 우리의 이해가 더욱 구체화될 수 있습니다.

이러한 고려 사항을 해결함으로써 향후 연구에서는 예측에서 증강 및 기타 모델 구성 요소의 역할에 대한 보다 포괄적인 이해를 제공하여 보다 효과적이고 해석 가능한 예측 모델을 위한 기반을 마련할 수 있습니다.

## 7 결론

저희는 정확한 CPI 예측을 위해 다변량 데이터와 감정 점수를 결합한 하이브리드 CNN-LSTM 모델을 사용하여 새로운 CPI 예측 프레임워크를 제안했습니다. 이전 연구와 비교했을 때, 제안된 모델은 눈에 띄는 성능 향상을 보여주었습니다. 하이브리드 CNN-LSTM 모델은 네 가지 핵심 기술을 통해 CPI 예측의 문제를 효과적으로 해결합니다. 첫째, 다변량 시계열 예측은 다양하고 복잡한 변수를 딥러닝 모델에 통합합니다. 둘째, 선형 보간을 통한 데이터 증강은 편향 없이 원래의 데이터 추세를 유지하면서 데이터 가용성이 제한된 경우를 완화합니다. 셋째, 뉴스 기사의 감정 분석은 코로나19, 정치, 자연재해와 같은 글로벌 이벤트의 비정형 데이터를 캡처하여 감정 점수로 변환합니다.

마지막으로 CNN-LSTM 하이브리드 접근 방식은 노이즈를 줄이고 시간적, 지역적, 글로벌 데이터 특성을 유지하는 CNN 계층을 통해 복잡한 CPI와 감정 지수에서 특징을 추출한 다음 LSTM 계층을 사용하여 정확한 CPI 예측을 수행합니다. 이 모델은 감정 분석을 통해 외부 이벤트를 인코딩하여 코호트와 같은 예기치 못한 상황에 대응할 수 있는 유연한 예측 프레임워크를 제공합니다.

## 작성자 기여

**개념화:** 서인재, 장백철.

**데이터 큐레이션:** 서인재.

**공식 분석:** 서인재.

**자금 조달:** 장백철.

**조사:** 서인재, .

**방법론:** 서인재, 김민경, 김종욱, 장백철.

**프로젝트 관리:** 장백철.

**리소스:** 장백철. **소프트웨어:** 서인재, . **감독:** 장백

철.

**Validation:** 서인재, 김민경, 김종욱, 장백철.

**시각화:** 서인재, .

**글쓰기 - 초안:** 서인재, .

**글쓰기 - 검토 및 편집:** 김종욱, 장백철.

## 참조

1. 바로 RJ 인플레이션과 성장. Rev 연방준비은행 세인트루이스 연은 1996;78:153-69
2. . 소비자 물가 지수; 2021. <http://kostat.go.kr/incomeNcpi/cpi/index.action>
3. 카발로 A. 코로나 소비 바구니로 인한 인플레이션. 2024;72(2):902-17. <https://doi.org/10.1057/s41308-023-00213-y>
4. Seiler P. COVID-19 시대의 가중치 편향과 인플레이션: 스위스 거래 데이터의 증거. 2020;156(1):1-11. <https://doi.org/10.1186/s41937-020-00057-7> PMID: 32959014
5. Jaravel X, O'Connell M. 기간 동안 인플레이션 급등과 제품 다양성 감소. 2020;191:104270. <https://doi.org/10.1016/j.jpubeco.2020.104270> PMID: 32952225

6. 디버트 WE, 폭스 KJ. 봉쇄 조건에서 실질 소비 및 CPI 편향 측정. NBER 워킹 페이퍼. 2020; (w27144).
7. 자하라 S. 딥러닝 알고리즘을 이용한 소비자물가지수에 대한 다변량 시계열 예측 기반 클라우드 컴퓨팅. In: 2020 제3회 정보 기술 및 지능형 시스템 연구에 관한 국제 세미나(ISRITI). IEEE; 2020. p. 338-43.
8. 드 올리베이라 카로시아 AE, 코엘류 GP, 다 실바 AEA. 브라질 주식 시장에 적용된 투자 전략: 딥러닝 감정 분석에 기반한 방법론. *Exp Syst Appl*. 2021;184.
9. 2017;73:136-40, ARIMA 모델에 기반한 Zhang X CPI 예측 *Adv Comput Sci Res*.
10. Ahmar AS 외. ARIMA 방법을 사용한 인도네시아의 CPI 예측 모델링 In: 제2회 통계, 수학, 교육 및 연구에 관한 국제 컨퍼런스 논문집. 2018.
11. 소비자 물가 지수 변화의 역학 분석 및 예측 Shinkarenko O, Hostryk A, Shynkarenko L, Dobrovol'ska I.. 동부-유럽 J 엔터 테크놀. 2021;113(109):40-51.
12. Liu Y. 결합된 ARIMA-LSTM 모델을 기반으로 중국의 소비자 물가 지수(CPI)를 예측합니다. *프론트 버스 경제 관리*. 2024;16(2):264-77.
13. 왕 A, 쉬 Y. 신장의 소비자 물가지수 마르코프 체인 분석 및 예측. 예: 2018 제 10 회 측정 기술 및 메카트로닉스 자동화 국제 컨퍼런스 (ICMTMA). 2018. p. 440-3.
14. Chou J-S, Ngo N-T. 건물 에너지 소비 패턴을 식별하기 위한 슬라이딩 윈도우 메타 휴리스틱 최적화 기반 머신 러닝 시스템을 사용한 시계열 분석. *Appl Energy*. 2016;177:751-70. <https://doi.org/10.1016/j.apenergy.2016.05.074>
15. 로마 MF, 푸트라 IKGD, 하르타티 RS, 아르디안토로 L. 가우스-방사선 기저 함수 커널로 동부 자바의 소비자 물가지수 도시 및 지역 예측. *J Phys*: 2020;1456(1):012026.
16. 모시리 S, 카메론 N. 인플레이션 예측에서 신경망 대 계량 경제학 모델. *J 예측*. 2000;19(3):201-17. [https://doi.org/10.1002\(sici\)1099-131x\(200004\)19:3<201::aid-for753>3.0.co;2-4](https://doi.org/10.1002(sici)1099-131x(200004)19:3<201::aid-for753>3.0.co;2-4)
17. 2024-T3 알루미늄 합금의 대기 부식 분석을 위한 신경망의 적용(Chen T, Yin M, Zhao Y, Wang S.). 2001;43(10):1931-50.
18. Szafranek K. 인플레이션 예측: 기계 학습 모델의 비교 분석 *J Risk Financ Manag*. 2019;12(4):174.
19. 맥아담 P, 맥넬리스 P. 두꺼운 모델과 신경망으로 인플레이션을 예측합니다. *Econ Model*. 2005;22(5):848-67. <https://doi.org/10.1016/j.econmod.2005.06.002>
20. 차우다리 MA, 하يدر A. 인플레이션 예측을 위한 신경망 모델: 평가. 2012;44(20):2631-5. <https://doi.org/10.1080/00036846.2011.566190>
21. 나카무라 E. 신경망을 이용한 인플레이션 예측. 2005;86(3):373-8. <https://doi.org/10.1016/j.econlet.2004.09.003>
22. 알모소바 A, 안드레센 N. 반복 신경망을 이용한 비선형 인플레이션 예측. ECB 워킹 페이퍼. 2019.
23. Zheng T, Wang Y, Wang D. 페널티가 적용된 MIDAS 회귀 및 뉴스 기반 감성을 이용한 CPI 예측. *Econ Model*. 2023;126:106186.
24. 계층적 순환 신경망을 이용한 CPI 인플레이션 구성 요소 예측. *Int J Forecast*. 2023;39(3):1145-62. <https://doi.org/10.1016/j.ijforecast.2022.04.009>
25. 호크라이터 S, 슈미트 후버 J. 장단기 기억. *Neural Comput*. 1997;9(8):1735-80. <https://doi.org/10.1162/neco.1997.9.8.1735> PMID: 9377276
26. 아키타 R, 요시하라 A, 마츠바라 T, 우에하라 K. 수치 및 텍스트 정보를 이용한 주식 예측을 위한 딥러닝. In: 2016 IEEE/ACIS 제15회 컴퓨터 및 정보 과학 국제 컨퍼런스(ICIS). 2016. p. 1-6. <https://doi.org/10.1109/icis.2016.7550882>
27. 성 엔, 남 케이. 예측을 위한 거시경제 효과와 감정 분석의 결합. *Entrue J Inf Technol*. 2017;16(2):41-54.
28. Li H, Xu Z, Taylor G, Studer C, Goldstein T. 신경망의 손실 풍경 시각화. In: *신경 정보 처리 시스템*; 2018.
29. 호세인자데 E, 하라티자데 S. CNNpred: 변수 세트를 사용한 CNN 기반 주식 시장 예측. 2019;129:273-85. <https://doi.org/10.1016/j.eswa.2019.03.029>
30. 리비에리스 IE, 핀텔라스 E, 핀텔라스 . 금 가격 시계열 예측을 위한 CNN-LSTM 모델. 2020;32(23):17351-60. <https://doi.org/10.1007/s00521-020-04867-x>
31. Ahmed A, Ali S, Raza A, Hussain I, Bilal A, Fitriyani NL 등. 단기 및 장기 에너지 소비를 동시에 예측하는 새로운 심층 신경망 아키텍처 융합. *PLoS ONE*. 2025;20(1):e0315668. <https://doi.org/10.1371/journal.pone.0315668> PMID: 39774457

32. 정 J, 장 B. 다변량 데이터 기반의 하이브리드 CNN-LSTM 모델을 이용한 정확한 전력 소비량 예측. PLoS ONE. 2022;17(11):e0278071. <https://doi.org/10.1371/journal.pone.0278071> PMID: 36417448
33. 트위터 정서를 이용한 코로나19 팬데믹에 대한 거시경제적 대응에 대한 국가 간 분석(Nia ZM, Ahmadi A, Bragazzi NL, Woldegerima WA, Mellado B, Wu J 등). PLoS ONE. 2022;17(8):e0272208. <https://doi.org/10.1371/journal.pone.0272208> PMID: 36001531
34. 송 Y, 윤 B. 소셜 미디어 데이터의 감정 분석을 이용한 감염병 예측. PLOS One. 2024;19(9):e0309842. <https://doi.org/10.1371/journal.pone.0309842> PMID: 39231189
35. 샤피로 AH, 수드호프 M, 윌슨 DJ. 뉴스 감정 측정. J 계량 경제학. 2020;228(2):221–43. <https://doi.org/10.1016/j.jeconom.2020.07.053>
36. Li J, Bu H, Wu J. 감정 인식 주식 시장 예측: 딥러닝 방법. In: 2017 서비스 시스템 및 서비스 관리에 관한 국제 컨퍼런스. 2017. p. 1–6. <https://doi.org/10.1109/icsssm.2017.7996306>
37. Li Y, Bu H, Li J, Wu J. 딥러닝의 향상을 통한 중국 주가 예측에서 텍스트 추출 투자자 감정의 역할. Int J Forecast. 2020;36(4):1541–62. <https://doi.org/10.1016/j.ijforecast.2020.05.001>
38. SKTBrain. KoBERT: 한국어 BERT 사전 훈련 케이스(KoBERT). 2019. <https://github.com/SKTBrain/KoBERT>
39. 박에스, 문제이, 김에스, 조위, 한제이, 박제이 외. Klue: 한국어 이해 평가. arXiv 사전 인쇄 2021. <https://arxiv.org/abs/2105.09680>
40. Teanaps. Teanaps: 텍스트 분석 API. 2021. 제공처: <https://github.com/fingeredman/teanaps#teanaps-text-analysis-apis>
41. In: 라스페이어 지수. 뉴욕, NY: 스프링거 뉴욕; 2008. p. 295-7. [https://doi.org/10.1007/978-0-387-32833-1\\_221](https://doi.org/10.1007/978-0-387-32833-1_221) PMID: 39247456
42. . 소비자 물가지수 산출 방법. 2025. <https://kostat.go.kr/menu.es?mid=b70101050000>
43. 통한 회귀 축소 및 선택. J Roy Statist Soc: Ser B (Methodol). 1996;58(1):267-88.
44. 미할시아 R, 타라우 P. TextRank: 텍스트에 질서를 부여하기. In: EMNLP 2004의 절차. 전산 언어학 협회; 2004.
45. Barrios F, Lopez F, Argerich L, Wachenchauser R. 자동 요약물을 위한 TextRank의 유사도 함수 변형. arXiv preprint 2016. <https://arxiv.org/abs/1602.03606>
46. 지브 M, 그레고르 A, 스토크백 FJ, 루이스 H, 멘도사 EM, 아마드니아 B. 추출 요약에서 체계적인 TextRank 최적화. In: 제14회 자연어 처리의 최근 발전에 관한 국제 컨퍼런스; 2023. 1274-83쪽.
47. 페이지 L, 브린 S, 모트와니 R, 위노그라드 T. 페이지랭크 인용 순위: 웹에 질서를 가져오다. 스탠포드 인포랩; 1999. p. 1999-66. 출처: <http://ilpubs.stanford.edu:8090/422/>
48. 말로 P, 신하 A, 코호넨 P, 왈레니우스 J, 타칼라 P. 좋은 부채 또는 나쁜 부채: 경제 텍스트에서 의미 지향성 감지. J Assoc Inf Sci Technol. 2014;65(4):782-96.
49. 이 E, 이 C, 안 S. 사전 학습된 언어 모델을 이용한 연구 제안서의 다중 클래스 텍스트 분류 비교 연구. Appl Sci. 2022;12(9):4522. <https://doi.org/10.3390/app12094522>
50. 앤트바일러 W, 프랭크 MZ. 모든 이야기는 소음일 뿐인가요? 인터넷 주식 게시판의 정보 콘텐츠. J Finance. 2004;59(3):1259-94.
51. 리 지아홍, 부 후이, 우 준지에. 감정 인식 주식 시장 예측: 딥러닝 방법. In: 2017 서비스 시스템 및 서비스 관리에 관한 국제 컨퍼런스. 2017. p. 1-6. <https://doi.org/10.1109/icsssm.2017.7996306>
52. 우 S, 리우 Y, 저우 Z, 왕 T-H. S-LSTM: 여러 데이터 소스 및 감정 분석에 기반한 주가 예측. 2022;34(1):44-62. <https://doi.org/10.1080/09540091.2021.1940101>
53. 차우 GC, 린 AL. 관련 계열에 의한 시계열의 최상의 선형 비편향 보간, 분포 및 외삽. Rev Econ Statist. 1971;53(4):372-5.
54. 어려운 시계열 데이터에 대한 신경망 예측을 개선하기 위한 프랙탈 보간 접근법, 라우비첵 S, 노이바우어 T. 2021;169:114474. <https://doi.org/10.1016/j.eswa.2020.114474>
55. 아키바 T, 사노 S, 야나세 T, 오타 T, 코야마 M. Optuna: 차세대 하이퍼파라미터 최적화 프레임워크. In: 제25회 지식 발견 및 데이터 마이닝에 관한 제25회 ACM SIGKDD 국제 컨퍼런스 프로시딩. 2019.
56. Kingma DP, Ba J. Adam: 확률적 최적화를 위한 방법 arXiv 프리프린트 2014. <https://arxiv.org/abs/1412.6980>

57. 드러커 H, 버지스 CJC, 카우프만 L, 스몰라 A, 바프닉 V. 자원 벡터 회귀 머신. In: 신경 정보 처리 시스템의 발전. 9 권 : MIT Press; 1997.
58. Breiman L. 무작위 숲. *Mach Learn.* 2001;45(1):5–32.  
<https://doi.org/10.1023/A:1010933404324>
59. 첸 T, 게스트린 C. XGBoost: 확장 가능한 트리 부스팅 시스템. In: 제22회 지식 발견 및 데이터 마이닝에 관한 제22회 ACM SIGKDD 국제 컨퍼런스 논문집; 2016. p. 785-94.
60. 정 J, 굴세레 C, 조 K, 벤지오 Y. 시퀀스 모델링에 대한 게이트 순환 신경망의 경험적 평가. *arXiv preprint* 2014.  
<https://arxiv.org/abs/1412.3555>
61. 바스와니 A, 샤제르 N, 파마르 N, 우즈코리트 J, 존스 L, 고메즈 AN 등. 주의만 기울이면 . In: 신경 정보 처리 시스템의 발전. Curran Associates, Inc.
62. 울슨 RF. 윌콕슨 사인-랭크 테스트. In: Wiley 임상시험 백과사전. Wiley; 2008. <https://doi.org/10.1002/9780471462422.eoct979>
63. 차이 T, 드렉슬러 RR. 평균제곱근오차(RMSE) 또는 평균절대오차(MAE)? - 문헌에서 RMSE를 피해야 한다는 주장에 대한 반론. *Geosci Model Dev.* 2014;7(3):1247–50. <https://doi.org/10.5194/gmd-7-1247-2014>
64. 김 티, 조 에스비. C-LSTM 신경망을 이용한 웹 트래픽 이상 탐지. 2018;106:66-76.