

시공간 의존성 네트워크 위상 및 그래프 신경망을 활용한 설명 가능한 환율 변화 예측 모형 개발

최인수¹, 고우성², 강기민¹, 장윤태², 노유진³, 이지윤³, 김우창*¹
¹ 한국과학기술원, ² 연세대학교, ³ 송실대학교

jl.cheivly@kaist.ac.kr; reiss.koh@yonsei.ac.kr; jytrewq1218@naver.com; kgm4752@kaist.ac.kr;
 n7348@naver.com; jy-0319@naver.com; wkim@kaist.ac.kr

Explainable Prediction Model of Exchange Rates via Spatiotemporal Network Topology and Graph Neural Networks

Insu Choi¹, Woosung Koh², Gimin Kang¹, Yuntae Jang², Yu Jin Roh³, Ji Yun Lee³, Woo Chang Kim¹

¹ Korea Advanced Institute of Science and Technology

² Yonsei University

³ Soongsil University

요 약

최근 환율 예측에 관한 다양한 연구가 진행되어 왔다. 이러한 추세에 대응하여 본 연구에서는 Pearson 상관 계수 및 상호 정보를 사용하여 외환 시장의 환율 변동을 분석하는 다중 연결 네트워크를 구축하였다. 본 연구에서는 이러한 구성된 환율 변화에 대한 시공간 의존성 네트워크를 만들고 그래프 기계 학습의 잠재력을 조사하여 예측 정확도를 향상시키려고 노력하였다. 본 연구 결과는 선형 및 비선형 종속 네트워크 모두에 대해 그래프 신경망을 활용한 임베딩을 활용하여 기존의 기계 학습 알고리즘과 결합시킬 경우 환율 변화의 예측력이 향상될 수 있음을 경험적으로 확인하였다. 특히, 이러한 결과는 통화 간 상호 의존성에만 의존하여 추가 데이터 없이 달성되었다. 이 접근 방식은 데이터 효율성을 강화하고 그래프 시각화를 통해 설명력 있는 통찰력을 제공하며 주어진 데이터 세트 내에서 효과적인 데이터를 생성하여 예측력을 높이는 결과로 해석할 수 있다.

1. 서론

높은 유동성과 지속적인 거래가 특징인 외환 (FOREX) 시장은 세계에서 가장 큰 금융 시장 중 하나로 알려져 있다(BIS, 2019). 환율 변동의 정확한 예측은 국제무역, 투자 결정, 리스크 관리 등에 영향을 미치기 때문에 투자자, 정책 입안자, 금융 기관 등 다양한 이해관계자에게 매우 중요하다(Frankel & Rose, 1995; Gosh et al., 2002). 따라서 많은 연구가 환율 변화를 예측하기 위한 효과적인 방법을 개발하기 위해 수행되어 왔다.

최근에는 전통적인 시계열 모델(Engle & Granger, 1987; Diebold et al., 1998)을 넘어서 기계 학습 알고리즘 기반 접근 방법(López et al., 2017)까지 환율 변화를 예

측하기 위한 다양한 시도가 이루어졌다. 본 연구에서는 외환 시장의 환율 변동을 분석하기 위해 피어슨 상관 계수와 상호 정보를 사용하여 시공간 네트워크를 구성하고 이를 예측에 활용하고자 한다.

본 연구에서는 환율 변동 예측을 위해 구성된 환율 변화에 대한 시공간 의존성 네트워크를 만들어 실제 환율 변화를 예측하는 데 있어 시공간 그래프를 활용한 그래프 신경망 방법론을 적용하였다(Battaglia et al., 2018; Scarselli et al., 2009). 그래프 기계 학습 기법은 다양한 영역에서 강력한 성능을 입증한 바 있기에 본 연구에서는 환율 시장에서의 환율 등락 예측을 위해 그래프 어텐션 네트워크(Graph Attention Network, GAT) (Veličković et al., 2017) 기법을 활용하였다.

2. 연구 방법론

2.1 데이터 설명

본 연구에서는 미국 달러(USD) 기준으로 가장 거래 빈번한 17개의 외환 시장의 환율 데이터를 활용하였다. 활용된 데이터는 다음과 같다: USD/JPY(일본 엔), USD/GBP(영국 파운드), USD/CNY(중국 위안), USD/AUD(호주 달러), USD/CAD(캐나다 달러), USD/CHF(스위스 프랑), USD/EUR(유로), USD/SEK(스웨덴 크로나), USD/NZD(뉴질랜드 달러), USD/SGD(싱가포르 달러), USD/HKD(홍콩 달러), USD/KRW(한국 원), USD/INR(인도 루피), USD/BRL(브라질 헤알), USD/ZAR(남아프리카 공화국 랜드), USD/TRY(터키 리라), USD/RUB(러시아 루블)이다.

본 연구에서는 60 일(3 개월)을 기준으로 이동창 간격을 1 일로 지정한 뒤 해당 기간 내의 Pearson 상관계수와 상호 정보량을 측정하였다. 각각의 이변량 상호 정보량을 계산할 때, 어떤 환율의 로그 변화량 데이터열에서도 7 개의 정규성 테스트를 만족하는 비율이 10% 이상을 넘어가는 데이터가 없었던 60 일 이동창 데이터가 활용되었다.

이러한 데이터들에 대해 60 일 (3 개월)을 기준으로 이동창 간격을 1 일로 지정한 뒤 해당 기간 내의 Pearson 상관계수와 상호 정보량을 측정하였다. 각각의 이변량 상호 정보량을 계산할 때, 하지만 본 연구에서 활용된 60 일 이동창 데이터는 어떠한 환율의 로그 변화량 데이터열에서도 7 개의 정규성 테스트를 만족하는 비율이 20% 이상을 넘어가는 데이터가 없었다. 즉, 해당 데이터가 정규성이 성립하지 않으며 이는 이변량 데이터에 대해 두 데이터가 이변량 정규분포를 만족할 경우 $I = -\frac{1}{2} \log(1 - \rho^2)$ 으로 동치인 관계가 성립하나(Ger'fand and Yaglom, 1957) 이러한 결과는 본 연구에서 Pearson 상관계수와 상호 정보량 모두를 사용하는 이론적 타당성을 뒷받침한다. 또한 네트워크의 연결을 위한 최소 p-value 임계치(threshold)는 0.1로 설정하였으며 상호 정보량은 $[0, 1]$ 의 범위를 갖도록 두 변수의 엔트로피의 최솟값을 분모로 갖는 정규화를 실시하였으며(Kvålseth, 2017), Markov 블록 부트스트랩(Markov block bootstrap)을 이용하여 p-value 를 계산하였다. 두 지표의 거리 함수에 대해서는 Pearson 상관계수의 경우에는 Euclidean 거리(Euclidean distance)를 활용한 식인 $\sqrt{2} \sqrt{1 - \rho}$ 를 사용하였으며, 상호정보량은 거리함수의 요건을 만족하는 $1 - \frac{MI(X,Y)}{\max(H(X),H(Y))}$ 를 활용하였다 (Kvålseth, 2017).

2.2 그래프 어텐션 네트워크를 활용한 임베딩

본 연구에서는 시공간 그래프 신경망 모형의 한 형태로 각 시점마다 본 Pearson 상관계수 그래프와 상호 정보량을 바탕으로 계산된 시공간 그래프를 바탕으로 그래프 어텐션 네트워크를 통해 학습하였다. 히스토그램을 이산화 방법으로 활용하였다. 각각의 거리는 거리 함수로 정의가 가능한 거리 지표를 활용하였다. 본 연구에서 제시하고자 하는 그래프 어텐션 네트워크는 임베딩 벡터의 형태로 활용되었다.

그래프 어텐션 네트워크 레이어는 다음과 같은 형태로 생성된다. N 을 그래프 내 정점의 수, F 를 정점에 속하는 특성의 수, F' 을 은닉 레이어의 길이라고 할 때 다음과 같은 정점의 특성(feature)이 존재하면

$$\mathbf{h} = \{\vec{h}_1, \dots, \vec{h}_N\}$$

위 \mathbf{h} 행렬은 그래프 어텐션 레이어를 통과하면 \mathbf{h}' 으로 변환되며 이 때 해당 행렬은 (N, F') 의 크기를 갖는다. $\mathbf{W} = (F', F)$, $\mathbf{a} = (2F', 1)$ 의 형태를 지닐 때 어텐션 계수는 다음과 같이 정의 된다.

$$e_{ij} = \mathbf{a}(\mathbf{W}\vec{h}_i, \mathbf{W}\vec{h}_j)$$

이 때 위 식은 정점 i 에 대한 정점 j 가 갖는 중요도를 의미한다.

이러한 어텐션 메커니즘은 단일 레이어 순방향 신경망으로 다음과 같은 학습 가능한 모수와 LeakyRELU 활성화 함수로 정의하게 된다. 이와 같이 계산된다.

$$\alpha_{ij} = \frac{\exp(\text{LeakyRELU}(\vec{a}^T[\mathbf{W}\vec{h}_i|\mathbf{W}\vec{h}_j]))}{\sum_{k \in N_i} \exp(\text{LeakyRELU}(\vec{a}^T[\mathbf{W}\vec{h}_i|\mathbf{W}\vec{h}_k]))}$$

이와 같이 계산된 어텐션 점수는 정점 i 에 대한 이웃의 중요도를 결정함으로써 투입 데이터를 새롭게 정의하게 된다.

$$\vec{h}'_i = \sigma \left(\sum_{k \in N_i} \alpha_{ij} \mathbf{W}\vec{h}_k \right)$$

이때, 특성으로 포함된 데이터는 각 FOREX 데이터의 20 일치 이동 평균과 그 전 5 일 시점까지로 그래프 특성을 포함하여 학습하였으며, 이는 원 모형 기반 실험에서는 데이터 열로 포함하여 실험을 실시하였다.

2.3 기계 학습 모형

본 연구에서는 기계 학습 모형의 다양성을 고려하여 예측 성능의 견고성을 향상시키기 위해 여러 가지

기계 학습 모델을 사용하였다. 이에 따라 로지스틱 회귀, 의사결정나무, XGBoost, LightGBM, CatBoost, 순환신경망(RNN), 장단기 메모리(LSTM), 게이트 순환 유닛(GRU) 및 Seq2seq 등의 모델을 활용하였다.

3. 실험 설명 및 결과

본 연구의 기간은 2015년 1월 1일부터 2022년 12월 31일이며, 학습 세트는 2015년 1월 1일부터 2020년 12월 31일까지로 설정되었다. 검증 기간은 2020년 6월 30일부터 2020년 12월 31일이며, 테스트 기간은 2021년 1월 1일부터 2022년 12월 31일까지다. 예측 문제는 분류 문제로 접근하여, 환율 상승과 하락에 각각 1과 0을 부여하고 이를 목표로 기계 학습 모델을 학습시켰다.

본 연구에서는 구성된 네트워크에 대해 GNN을 활용한 임베딩을 수행하였고, 이를 다양한 기계 학습 모델의 기본 모델과 시공간 그래프 신경망을 사용한 벡터 모델로 특성 데이터를 추출하여 예측에 활용하였다. 기초 모델과 GNN과 결합한 모델 모두 초모수 최적화를 수행하여 17개 예측 결과에 대한 평균 예측 정확도에서 Pearson 상관계수와 상호 정보량 네트워크를 활용한 모델의 정확도가 전체 평균적으로 약 1%에서 3% 가량 향상되었으며, 상호 정보량 네트워크를 활용한 결과가 환율 등락 예측에서 Pearson 상관계수를 활용한 네트워크보다 조금 더 나은 성능을 보임을 확인하였다.

4. 결론

본 연구 결과로 인해, Pearson 상관계수와 상호 정보량을 기반으로 한 선형 및 비선형 종속 네트워크 모두에 대해 그래프 신경망을 활용한 임베딩을 기존 기계 학습 알고리즘과 결합할 경우 환율 변화의 예측력이 향상된 것을 확인할 수 있다. 이러한 결과는 통화 간 상호 의존성에만 의존하여 추가 데이터 없이도 달성되었다. 이 방법은 데이터 효율성을 강화하고 그래프 시각화를 통한 설명력을 제공하며, 주어진 데이터 세트 내에서 효과적인 데이터를 생성하여 예측력을 높이는 결과로 해석할 수 있다. 이러한 개선된 예측력은 금융 기관과 개인 거래자들이 거래 전략을 개선하고, 리스크 관리를 향상시키며, 잠재적으로 투자 수익률을 제고하는 기회를 발견할 수 있도록 도움을 줄 것으로 기대된다.

또한, 이 연구에서 제안된 접근법은 다양한 금융 데이터를 효과적으로 분석하고 이해하는 데 도움이 되며, 이를 통해 금융 시장의 안정성과 효율성을 높일 수 있다. 향후 연구에서는 더 많은 금융 데이터와 다양한 종류의 기계 학습 모델을 적용하여 연구의 범위를 확장하고, 그래프 신경망 및 기타 기계 학습 기법의 성능을 더욱 향상시킬 수 있는 방안을 모색할 예정이다. 이렇게 함으로써, 금융 시장 예측의 정확성과 효율성을 더욱 높이는 데 기여할 수 있을 것으로 기대된다.

참고문헌

- [1] Bank for International Settlements. (2019). Triennial Central Bank Survey: Foreign exchange turnover in April 2019. Retrieved from https://www.bis.org/statistics/rpfx19_fx.pdf
- [2] Frankel, J. A., & Rose, A. K. (1995). Empirical research on nominal exchange rates. *Handbook of international economics*, 3, 1689-1729.
- [3] Ghosh, A. R., Gulde, A. M., & Wolf, H. C. (2002). *Exchange rate regimes: choices and consequences*. MIT press.
- [4] Diebold, F. X., Husted, S., & Rush, M. (1991). Real exchange rates under the gold standard. *Journal of Political Economy*, 99(6), 1252-1271.
- [5] Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2), 251-276.
- [6] López, J. P., Cruz, G., & Sandoval, J. (2017). A deep learning method for short-term forecasting of the EUR/USD exchange rate. *MATEC Web of Conferences*, 137, 09003.
- [7] Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., ... & Lillicrap, T. (2018). Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*.
- [8] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1), 61-80.
- [9] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. *arXiv preprint arXiv:1710.10903*.
- [10] Gel'fand, I. M., & Yaglom, A. M. (1957). Computation of the amount of information about a stochastic function contained in another such function. *Uspekhi Matematicheskikh Nauk*, 12(1), 3-52.
- [11] Kvålseth, T. O. (2017). On normalized mutual information: measure derivations and properties. *Entropy*, 19(11), 631.