

감정분석과 GRU 모델 기반의 주가지수 예측 및 분석: 코스피 200 에너지/화학을 중심으로

양혁진*¹, 최종민*², 박찬정**

Stock Index Prediction and Analysis based on Sentiment Analysis and GRU Models: Focused on KOSPI 200 Energy/Chemicals

Hyeok-Jin Yang*¹, Jong-Min Choi*², and Chan Jung Park**

이 논문은 2024학년도 제주대학교 교원성과지원사업에 의하여 연구되었음

요 약

코로나19 팬데믹 이후, 주식 시장의 변동성의 폭이 증가하였다. 과거에는 재무 분석과 통계적 방법을 이용하여 주가지수를 예측하려는 경향이 많았으나, 최근에는 인공지능 기술을 기반으로 하여 주가지수를 예측하려는 시도가 이루어지고 있다. 본 연구는 코스피 200 에너지/화학 주가를 중심으로 개인 투자자를 위해 감정분석과 GRU와 LSTM 딥러닝 모델 기반의 주가지수를 각각 예측한 후, 성능 평가를 진행한다. 이를 위해 뉴스 데이터를 크롤링하였고 Naver Clova API와 KR-FinBERT-SC를 활용하여 감정분석을 수행하였으며, 예측의 정확도를 높이기 위해 산출된 감정 점수와 주가지수를 결합하여 딥러닝 모델에 각각 적용하였다. 그 결과, GRU 모델이 LSTM 모델보다 코스피 200 에너지/화학 주가지수 예측에 더 적합함을 확인하였고, 비록 한정된 기간의 데이터를 사용했지만, 감정분석과 GRU 딥러닝 알고리즘의 결합으로 예측 정확도를 높였다.

Abstract

Since the COVID-19 pandemic, the volatility of the stock market has increased. In the past, there was a tendency to predict stock indices using financial analysis and statistical methods, but recently, attempts have been made to predict stock indices based on artificial intelligence technology. This paper focuses on predicting the KOSPI 200 energy/chemical stock index for individual investors using sentiment analysis and GRU and LSTM deep learning models, and then evaluates their performance. To achieve this, news data was crawled and sentiment analysis was performed using Naver Clova API and KR-FinBERT-SC. The derived sentiment scores were combined with the stock index and applied to each deep learning model to improve prediction accuracy. As a result, it was found that the GRU model is more suitable than the LSTM model for predicting the KOSPI 200 energy/chemical stock index. Despite using data from a limited period, the combination of sentiment analysis and the GRU deep learning algorithm improved prediction accuracy.

Keywords

stock price prediction, sentiment analysis, KR-FinBERT-SC, GRU, individual investors

* 제주대학교 컴퓨터교육과 학사과정
- ORCID¹: <https://orcid.org/0009-0007-6486-8174>
- ORCID²: <https://orcid.org/0009-0002-1465-5924>
** 제주대학교 컴퓨터교육과 교수(교신저자)
- ORCID: <https://orcid.org/0000-0002-3146-2342>

• Received: Mar. 17, 2024, Revised: Jun. 08, 2024, Accepted: Jun. 11, 2024
• Corresponding Author: Chan Jung Park
Dept. of Computer Education, Jeju National University, Jejudachak-ro 102
63243, South Korea
Tel.: +82-64-654-3296, Email: cjpark@jejunu.ac.kr

1. 서 론

전 세계를 강타한 코로나19 팬데믹과 함께 경제적 변동성이 증대됨에 따라, 주식 시장에 대한 개인 투자자의 관심이 급격히 증가하였다. 그러나 많은 개인 투자자가 시장의 변동성을 제대로 이해하지 못하고, 높은 투자 위험에 자신의 자산을 노출하고 있는 것으로 나타났다[1]. 이는 개인 투자자가 주식 투자의 중요성에 비추어 볼 때, 위험관리에 대한 인식이 부족하고, 실질적인 위험관리 전략을 수립하고 실천하는 과정에서 어려움을 겪고 있기 때문이다[1]. 또한, 개인 투자자가 시장 정보를 분석하고 자신의 포트폴리오를 체계적으로 관리하는 데 필요한 시간과 기술력이 부족하다는 점도 큰 문제로 지적되고 있다[2]. 기업가치의 불확실성이 높고 주가 수익률 변동성이 큰 중소형주 또는 의료 섹터에 대한 비중이 높은 가운데, 소수의 종목에 집중적으로 투자하고 있어 개별주식 고유위험(Idiosyncratic risk)의 분산 효과를 기대하기 어렵기 때문이다. 고유위험이란 회사의 경영 상태, 산업 내 경쟁, 경영진의 결정, 제품의 성공 여부 등 특정 기업에만 영향을 미치는 요인들에 의해 발생한다. 이런 고유위험은 주식의 포트폴리오화(Portfolio)로 감소시킬 수 있다. 하지만, 포트폴리오 위험이 발생할 수 있으며, 이때, 포트폴리오 위험이란 포트폴리오에 포함된 여러 주식의 전체적인 위험도를 의미하고, 주식 시장의 변동성과 개별주식의 특성에 영향을 받게 된다[2].

주식 포트폴리오 위험성은 기존투자자보다는 신규투자자, 여성보다는 남성, 고액 투자자보다는 소액투자자가 더 높은 것으로 파악된다[2]. 이는 개인 투자자가 분산투자 및 위험관리에 대한 중요성을 제대로 인지하지 못하는 경우도 있지만, 시간적·기술적인 한계 탓에 위험자산 포트폴리오를 체계적으로 관리하기 어렵기 때문이다. 따라서 투자 위험 및 기업에 대한 사전적 분석 없이 단기적인 고수익을 노리는 주식매매가 자주 발생하곤 한다[1]. 개인 투자자가 모든 정보를 분석해 주가를 예측하는 건 불가능하며, 주어진 정보가 긍정적일지 부정적일지를 파악하는 일은 상황에 따라 개인차가 많다.

최근까지 주가 예측 연구는 수치 지표를 사용해

예측 모델을 만들어왔다[3]-[5]. 예를 들어, BLSTM(Bidirectional Long Short-Term Memory)을 이용한 주가지수 예측 연구는 기존의 LSTM(Long Short-Term Memory) 방식의 단방향 예측에서 발생하는 오류를 줄이고, 주식 가격에 영향을 미치는 거시 지표를 분석하여 주가지수를 예측하였다[3]. 야후에서 제공하는 yfinance 주가 데이터의 경우에는 GRU(Gated Recurrent Unit)의 성능이 LSTM에 비해 좋았다[6]. GRU가 LSTM보다 학습할 가중치가 적은 장점이 있어 상황에 맞는 LSTM과 GRU에 대한 비교 및 분석이 필요해 보인다[6].

한편, 주가 예측에서 뉴스 기사에는 다양한 의견과 해석이 포함되어 있고 실제 의미를 중립적으로 전달되어 파악하기 힘들어 그동안 연구에서 뉴스를 활용한 사례는 많지 않다[7]. 하지만 뉴스가 전달하는 정보는 투자자들의 기대와 판단에 영향을 미쳐 결과적으로 주가 변동을 일으킨다[8]. 또한, 뉴스는 시계열 속성을 갖는다. 시계열 데이터 예측은 경제학, 금융학뿐만 아니라 다양한 분야에서 중요한 연구 주제이다. 특히, 주가지수 예측은 불확실성이 높은 시장 환경에서 매우 좋은 효과를 보여주었다[8].

본 연구는 주식 시장의 실제 매매 데이터를 활용하여 주가 변동성을 정확하게 예측하고자 한다. 기존 연구들이 주로 긍정적 혹은 부정적인 의견만을 바탕으로 주가의 흐름을 단순화하여 설명하는 한계를 넘어서, 본 연구에서는 증권사 리포트, 재무제표 등과 같은 정형 데이터와 SNS의 감정분석 데이터, 뉴스 기사의 텍스트와 같은 비정형 데이터를 결합하여 복잡한 시장의 동향을 더욱 정확하게 분석하고자 한다.

기존 연구에서 주가 예측과 관련한 실험을 진행하였을 때, S&P 500, KOSPI 200과 같이 보편적으로 널리 사용되는 투자 지수를 활용하여 다양한 종목들에 대해 실험을 수행하였을 때 효과적이었다[9]. 본 연구는 코스피 200지수 중에서도 특정 기업의 영향력이 지나치게 크지 않으며, 상대적으로 다양한 산업군이 균형 있게 포함된 에너지/화학 지수를 예측 대상으로 선정하였다. 이를 통해, 시장의 전반적인 트렌드를 파악하고, 개별 기업의 영향력이 과도하게 반영되지 않는 객관적인 분석을 수행하고자 한다.

또한, 본 연구에서는 뉴스와 같이 시계열 속성을 가진 텍스트 데이터로부터 긍정과 부정의 감정분석을 수행하고 이를 반영하여 GRU 모델과 LSTM 모델로 주식 데이터를 학습시켜, 주가를 더욱 정확히 예측하기 위한 프로그램을 개발하였다. 이때, 주식 데이터는 KRX 코스피 200 에너지/화학 데이터를 웹 크롤링을 통해 구축하였고, NAVER CLOVA Studio와 KR-FinBERT-SC를 사용하여 감정 점수를 구하였다. 주요 성능 평가 지표인 평균 절대 오차 (Mean absolute error)와 평균 제곱 오차의 제곱근 (Root mean square error)을 통해 본 연구에서 진행한 감정분석과 딥러닝 기반의 주가지수 예측에 있어 GRU 모델 기반 주가 예측이 LSTM 모델 기반 보다 좋은 성능을 가짐을 확인하였다.

II. 관련 연구

2.1 KR-FinBERT

BERT(Bidirectional Encoder Representations from Transformers)는 텍스트 데이터의 깊은 분석과 정확도로 주식 가격 예측에 자주 쓰인다[10]-[12]. 경제 뉴스, 기업 보고서, 소셜 미디어 같은 다양한 데이터가 주식 시장에 큰 영향을 주며, BERT는 이 데이터 속 복잡한 패턴과 관계를 이해해 가격 변동을 예측하는데 필요한 통찰을 제공한다. 특정 도메인 데이터로 미세 조정이 가능해 금융 분야에서 유용하며, 복잡한 의미 관계와 패턴 인식은 변동성이 큰 주식 시장에서 유의미한 예측을 하는 데 도움을 준다[10]-[12].

KR-FinBERT는 한국어 특화 모형 KR-BERT를 기반으로 하며, 금융 텍스트 데이터로 학습된 언어 모형이다. KR-BERT는 작은 크기의 한국어 말뭉치를 사용해 학습 시간을 단축했으며, 기존 모델들이 필요로 하는 많은 시간과 비용을 줄였다[13][14]. 이 모델은 위키피디아 한국어 데이터, 뉴스 기사, 법률 정보, 댓글 등 일반 텍스트와 경제/금융 뉴스, 증권사 리포트 등 금융 전문 데이터를 포함한 대규모 데이터셋으로 훈련되었다[14]. 훈련 데이터셋에는 440,067개의 뉴스 제목과 11,237개의 보고서 문서가 있으며, 총 크기는 약 13.22GB에 달한다[14].

모델은 데이터의 각 문장을 독립적인 입력으로

취급해 6,379,315개의 입력 데이터로부터 학습했다. 학습 과정에서는 최대 문장 길이를 512, 배치 크기를 32, 학습률을 $5e-5$ 로 설정하고, GPU를 사용해 총 5천만 스텝 이상을 수행했다. 이러한 방식으로 KR-FinBERT는 금융 분야의 복잡한 언어 패턴을 이해하고 처리하는 데 최적화되었다[14]. 본 연구에서는 정형 데이터와 크롤링한 비정형 데이터를 요약한 후, KR-FinBERT-SC를 기반으로 감정분석을 실행하였다. KR-FinBERT-SC는 데이터 처리시 데이터 양에 대한 제한이 있어, 모델의 입력에 알맞게 데이터를 요약한 후, 감정분석을 진행한다. 데이터 요약은 NAVER CLOVA Summary API[15]를 사용하였다.

2.2 감정분석과 딥러닝 알고리즘을 함께 고려한 기존 연구

주가 예측은 과거부터 많은 연구자의 관심을 받아온 복잡한 문제였다. 그러나 최근에는 다양한 방법론과 접근법이 제안되며 이 문제를 해결하기 위한 시도가 활발히 이루어지고 있다[11][16][17].

기존 연구는 뉴스나 거시경제 지표를 단독 혹은 제한적으로 사용했다. 장은아 등의 연구는 미국 다우존스지수를 예측하기 위해 뉴스의 감정분석과 다양한 거시경제지표를 결합한 새로운 접근을 시도했다[11]. 장주현과 김재운의 연구에서는 뉴스 데이터를 활용하여 뉴스의 감성을 반영해 주가지수를 예측하는데 XGBoost와 감정분석을 했고, 자연어 처리를 위해서는 한국 금융에 맞는 KR-FinBERT를 사용했다[16]. 이 연구에서는 뉴스의 전처리된 뉴스의 제목으로 긍정, 중립, 부정의 비율을 환산해주어 뉴스가 긍정인지 부정인지 파악한 후, 수치 지표와 함께 주가지수를 예측하는데 사용하였다. 하지만, 제목만을 통해 감정분석을 하였고, 뉴스 본문 전체를 대상으로 하지는 않아 제한점이 있다[16].

한편, 공모주 투자에 대한 개인 투자자들의 관심과 참여가 증가하고 있으며, 이에 따른 영향력과 투자 수익에 관한 연구가 확대되고 있어, 개인 투자자의 관심이 공모주 시초가에 영향을 준다고 가정하고, 최남현 등의 연구에서는 뉴스와 댓글을 통한 감정분석을 포함한 다양한 변수를 사용하여 공모주 가격 형성을 분석했다[17].

XGBoost와 랜덤 포레스트 모델을 사용했으며, XGBoost가 더 좋은 성능을 보였다. 감성 변수 포함 시 초과 수익 예측에는 도움이 되지만, 초과 상승분을 예측하는 것은 어려움을 보였다. 연구는 인터넷 검색 횟수를 넘어서 댓글과 뉴스를 통해 개인의 관심도를 감성 변수화하여 분석함으로써 의의가 있다. 하지만 기관투자자의 관심도, 다양한 데이터 소스의 부족, 공모가액 규모별 분석의 필요성 등 한계도 지적됐다[17].

감정분석에는 규칙과 어휘 기반의 VADER(Valence Aware Dictionary and sentiment Reasoner), 예측 모델에는 LSTM을 사용해 효과적인 지표 조합을 찾았다. 본 연구에서는 사전 기반의 감정분석 대신, 오랫동안 축적된 방대한 데이터를 기반으로 텍스트에 담긴 감정분석 방법을 위주로 하여 뉴스를 기반으로 좀 더 기업에 대한 정확한 긍정과 부정의 감정을 도출하고자 하였다[18].

III. 주가지수 예측 프로그램 설계

3.1 GRU 모델과 LSTM 모델

GRU는 순환 신경망(RNN)의 한 형태로, LSTM보다 더 간단한 구조를 가지고 있다. GRU는 셀 상태와 은닉 상태를 하나로 통합하여 가중치 학습의 복잡성을 줄이고 계산 효율성을 높인다[6][19]. 이러한 특징 덕분에 GRU는 많은 시계열 예측 문제에서 LSTM보다 더 나은 성능을 보일 수 있다.

LSTM은 순환 신경망의 한 형태로, 장기 의존성 문제를 해결하기 위해 설계되었다. LSTM은 셀 상태를 유지하며 게이트 구조(입력 게이트, 망각 게이트, 출력 게이트)를 통해 정보를 선택적으로 기억하고 잇는다. 이로 인해 LSTM은 긴 시퀀스 데이터를 처리하는 데 강력한 성능을 발휘한다[3][20][21].

주가 예측에서 두 모델을 적용한 사례가 많았으며, 어떤 상황에서 어떤 모델이 더욱 좋은 성능을 나타내는지 두 모델에 관한 분석 연구가 필요하였고, 본 연구와 같이 제한된 데이터를 가진 상황에서 GRU 모델의 적절성을 제시한 연구[6]를 바탕으로 주가 예측의 결과를 분석하였다.

3.2 자료 수집

행동경제학과 비정형 데이터 마이닝은 최근 자산 시장 연구에서 중요한 이슈로 부상하고 있다. 특히, 투자자의 심리를 이해하는 데 있어 비정형 데이터 마이닝은 기존의 접근법에 대한 새로운 시각을 제공한다. 본 연구에서는 정형, 비정형 데이터의 감정분석 값을 활용하여 주가지수를 예측하는 방법론을 제시한다. 이러한 연구 결과들은 감정분석이 주가 예측에 큰 영향을 미칠 수 있음을 입증하였으며[16], 본 연구에서는 정형, 비정형 데이터의 감정분석 값을 활용하여 주가를 예측하였다.

3.2.1 비정형 데이터

본 연구에서 사용할 비정형 데이터는 2023년 1월 1일부터 2023년 9월 30일까지의 기간 동안 네이버 뉴스에서 수집된 에너지/화학 관련 기사이다. 데이터 수집을 위한 기간 설정의 근거는 최근의 경제 정세와 시장 상황을 반영하여 예측의 정확성을 높이기 위함이다. 과거의 데이터를 너무 많이 반영하게 되면 현재의 시장 상황과 일치하지 않는 예측 결과를 초래할 수 있다. 데이터 수집 과정에서는 웹 크롤링 도구를 활용하여 데이터를 자동 수집하였다. 이를 위해 그림 1과 같이 크롤링 프로그램을 설계하였다.

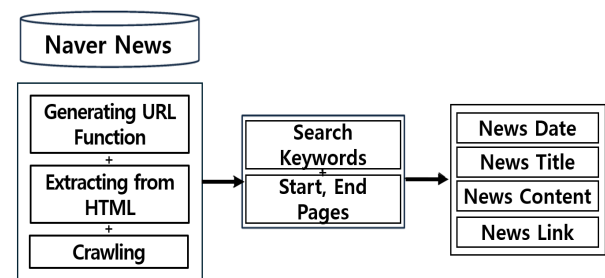


그림 1. 네이버 뉴스 데이터 크롤링

Fig. 1. Naver news data crawling

3.2.2 정형 데이터

본 연구에서 사용할 정형 데이터는 2023년 1월 1일부터 2023년 9월 30일까지의 기간 동안 발표된 증권사 리포트이다.

이 데이터는 주가 예측에 중요한 거시적 지표들을 포함하고 있다. 따라서 유가, 국제 금값, 환율 등의 거시적 지표는 증권사 리포트와 중복되는 정보를 포함할 수 있으므로 별도로 수집하지 않았다. 정형 데이터의 수집 기간 역시 최근의 경제 정세와 시장 상황을 반영하기 위해 설정되었다. 과거의 데이터를 너무 많이 반영하게 되면 현재의 시장 상황과 일치하지 않는 예측 결과를 초래할 수 있다.

3.3 감정분석

CLOVA Studio는 NAVER사에서 개발한 대규모 전문 데이터와 AI 모델을 결합하여 기업 맞춤형 AI 서비스를 제작할 수 있는 개발 도구이다[4]. 기업이 보유한 도메인별 빅데이터를 HyperCLOVA X 언어 모델과 접목시켜 비즈니스 요건에 특화된 AI 애플리케이션을 효율적으로 구축할 수 있어 이를 통한 맞춤형 AI 서비스 개발로 기업의 업무 효율성 제고와 경쟁력 강화에 기여할 수 있다[15]. 또한, 기업 현업에서 발생하는 다양하고 복잡한 문제들을 AI의 힘을 빌려 효과적으로 해결할 수 있다[15]. 본 연구에서는 그림 2와 감정점수를 추출하였다. 정형 데이터도 같은 과정을 통해 감정분석이 진행되었다.

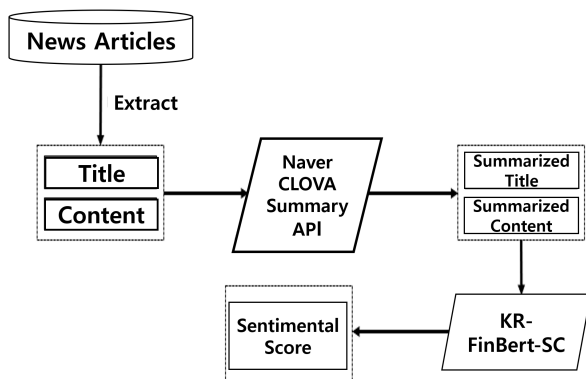


그림 2. 감정분석 과정

Fig. 2. Process for sentiment analysis

3.4 주가-감정 통합 지수 산출법

본 연구에서는 세 가지 요소를 결합하여 결과값을 도출하는 방법을 그림 3과 같이 제시한다.

$$(\text{실제 지수} \times \text{지수 비중}) + (\text{비정형 데이터 비중} \times \text{비정형 데이터 감정 점수}) + (\text{정형 데이터 비중} \times \text{정형 데이터 감정 점수})$$

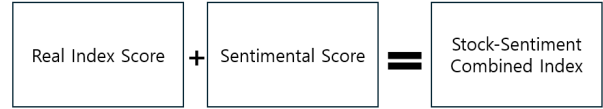


그림 3. 주가-감정 통합 지수 산출법

Fig. 3. Stock-sentiment combined index calculation method

(실제 지수 × 지수 비중) 첫번째 요소는 ‘실제 지수’와 ‘지수비중’의 곱이다. ‘실제 지수’는 KRX 코스피 200 에너지/화학의 지수이며, ‘지수 비중’은 해당 실제 지수가 전체 결과값에 미치는 상대적인 중요도를 나타낸다.

(비정형 데이터 비중 × 비정형 데이터 감정 점수) 두번째 요소는 ‘비정형 데이터 비중’과 ‘비정형 데이터 감정 점수’의 곱이다. ‘비정형 데이터 감정 점수’는 비정형 데이터에서 추출한 감정 정보를 점수화한 것이며, ‘비정형 데이터 비중’은 해당 데이터가 전체 결과값에 미치는 상대적인 중요도를 나타낸다.

(정형 데이터 비중 × 정형 데이터 감정 점수) 세번째 요소는 ‘정형 데이터 비중’과 ‘정형 데이터 감정 점수’의 곱이다. ‘정형 데이터 감정 점수’는 정형 데이터에서 추출한 감정 정보를 점수화한 것이며, ‘정형 데이터 비중’은 해당 데이터가 전체 결과값에 미치는 상대적인 중요도를 나타낸다. 이 공식을 통해 각 데이터의 비중을 고려하여 종합적인 예측 결과를 도출한다. 이렇게 도출된 결과는 주가지수의 예측에 사용된다.

3.5 평가 지표

모델 예측값의 정확성에 대한 평가 지표는 다음과 같다[22]. MAE(mean absolute error)는 식 (1)과 같이 실제값과 예측값의 차이를 절댓값으로 변환한 평균이다. 이 값이 작을수록 예측이 더 정확하다는 것을 의미한다.

$$MAE = \frac{1}{n} \sum |\hat{y} - y| \quad (1)$$

RMSE(Root Mean Squared Error)는 식 (2)와 같이 실제값과 예측값의 차이를 제곱해서 평균한 값에 루트를 취한 값이다.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum(\hat{y}-y)^2}{n}} \quad (2)$$

RMSLE(Root Mean Squared Logarithmic Error)는 식 (3)과 같이 MSLE의 제곱근을 구한 값이다. MSLE는 실제 값과 예측값에 각각 로그를 취한 후, 그 차이를 제곱하여 평균을 낸 값을 의미한다. RMSLE는 실제 값과 예측값의 비율에 더 초점을 맞추며, 이 값이 작을수록 예측이 더 정확하다는 것을 의미한다.

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(\hat{y}_i + 1) - \log(y_i + 1))^2} \quad (3)$$

R²는 결정계수로 회귀 모델의 설명력을 의미한다. 식 (4)와 같이 실제 값의 분산을 모델이 얼마나 잘 설명하는지를 나타내는 지표로, 1에 가까울수록 모델이 데이터를 잘 설명하고 있다는 것을 의미한다. 반면, 0에 가까울수록 모델이 데이터를 잘 설명하지 못하고 있다는 것을 의미한다.

$$R^2 = \frac{SS_{trt}}{SST} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (4)$$

IV. 구현 사항 및 분석

4.1 데이터 수집 및 분석

본 연구에서는 BeautifulSoup, 파이썬의 판다스(Pandas) 라이브러리를 활용하여 2023년 01월 01일부터 2023년 9월 30일까지 코스피 200 에너지/화학 기업 관련 1,014개의 데이터를 수집하였다. 또한, 같은 기간 동안 코스피 200 에너지/화학 지수에 포함된 기업의 증권사 리포트 총 32개를 그림 4와 같이 수집하였다.

HD현대일렉트릭	2023-10-17 오후 7:27	파일 불러
LG화학	2023-10-17 오후 7:27	파일 불러
OCI홀딩스	2023-10-18 오후 5:04	파일 불러
Pi점단소재	2023-10-17 오후 7:27	파일 불러
SKC	2023-10-17 오후 7:27	파일 불러
SK아이이테크놀로지	2023-10-17 오후 7:27	파일 불러
SK이노베이션	2023-10-17 오후 7:27	파일 불러
S-Oil	2023-10-17 오후 7:27	파일 불러
금호석유	2023-10-17 오후 7:27	파일 불러
롯데케미칼	2023-10-17 오후 7:27	파일 불러
한솔케미칼	2023-10-17 오후 7:27	파일 불러
한화	2023-10-17 오후 7:27	파일 불러
한화솔루션	2023-10-17 오후 7:27	파일 불러
후성	2023-10-17 오후 7:27	파일 불러

그림 4. 증권사 리포트

Fig. 4. Securities company report

수집한 정형 데이터와 비정형 데이터를 그림 5와 같이 Naver CLOVA Summary API를 사용하여 요약하였고 이후 KR-FinBERT-SC 사용에 적합한 자료 형태로 변환하였다.

content
{유상증자 발표 후 두 달만에 추가 33.8% 급락 배터리 사업 확장에 재무 부담....}
{SK그룹은 지주사 체제 전환 이후 동일인(총수)과 총수일가의 낮은 지배력을 보완하기....}
일반 공모 청약 경쟁률 67.8대 1 기록했지만 기간과 경쟁률은 14.89대 1에 그쳐....}
{\n제7회 GMF 본선 무대 열려...6개팀 경연대상 1000만원 등 총 2100만원....}
{\n지배구조 정점 SK... 최태원 지분율 17.5%자녀들 지분은 '0'... 상속시....}

(a) 요약 전 데이터

(a) Data before summary

No	text
1	SK그룹은 지주사 체제 전환 이후 동일인(총수)과 총수일가의 낮은 지배력을 보완하기....
2	CGV보다 낮아상장일은 멀고 2차전지 추가 하락세 "분위기에 휩쓸려 청약하긴 했는데....
4	최태원 지분율 17.5%자녀들 지분은 '0'... 상속시 지분율 확충이 최태원 "자식이....
5	왼쪽이 조지아 2공장, 오른쪽이 조지아 1공장 SK이노베이션의 추가 발목을 잡았던 미....
6	SK이노베이션은 자원개발 자회사인 SK아스온이 남중국해에 있는 해상 유전에서 이달부....
7	올해로 7회째...파라솔 클라리넷 앙상블 팀, 대상문체부 장관상 SK이노베이션이 후....

(b) 요약 후 데이터

(b) Data after summary

그림 5. 데이터 요약 전/후

Fig. 5. Before/After data summary

이후, KR-FinBERT-SC를 사용하여 정형, 비정형 데이터의 감정분석을 실시하였다. 감정분석 결과는 그림 6과 같이 긍정(Positive), 부정(Negative), 중립(Neutral)으로 표현되며, 이 각각은 0과 1 사이의 값을 가진다.

text	sentiment	score
SK그룹은 지주사 체제 전환 이후 동일인(총수)과 총수일가의 낮은 지배력을 보완하기...	netural	0.977278
CGV보다 낮아상장일은 멀고 2차전지 주가 하락세 '분위기에 휩쓸려 청약하긴 했는데...	Negative	0.989004
최태원 지분을 17.5%자녀들 지분은 0... 상속시 지분을 확충여최태원 '자식이...	netural	0.999888
왼쪽이 조지아 2공장, 오른쪽이 조지아 1공장SK이노베이션의 주가 발목을 잡았던 미...	netural	0.888966
SK이노베이션은 자원개발 자회사인 SK머스온이 남중국해에 있는 해상 유전에서이달부...	positive	0.999302
올해로 7회째...파라솔 클리넷 양상불 팀, 대상문제부 장관상SK이노베이션이 ...	netural	0.552830

그림 6. 감정분석 결과

Fig. 6. Results of sentiment analysis

분석 결과값을 주가지수와 결합하는 과정에서, (-1~+1)의 범위에서 부정적인 감정은 음수로, 긍정적인 감정은 양수로 치환하였다. 중립적인 감정은 글의 표현이 중립적이거나 주가의 변동이 크지 않을 수 있음을 의미한다. 따라서 중립적인 감정은 주가의 상승이나 하락에 직접적인 영향을 주지 않는다는 판단하에 계산에서 제외하였다. 이렇게 설정한 이유는 감정의 극성에 따라 주가의 변동성을 보다 정확히 반영하기 위함이다. 또한, 각각의 데이터에 대한 비중을 적용하는 과정에서, 주가지수의 비중을 0.6, 비정형 데이터의 비중을 0.05, 정형 데이터의 비중을 0.35로 설정하여 통합하였는데, 이 비율 조합에서 가장 높은 예측값이 도출되어 이를 바탕으로 해당 비중을 최종 모델에 적용하였다.

4.2 예측 모델

본 연구에서는 감정분석과 GRU 모델과 LSTM 모델을 활용하여 주가 예측 모델을 개발하였다. 주식 데이터는 시계열 데이터이므로, DNN과 CNN에 비해 시계열 예측에 적합한 LSTM과 BiLSTM 모델의 성능이 월등히 높음을 알 수 있다[23]. 따라서 본 연구에서도 시계열 예측에 적합한 LSTM 모델과 GRU 모델과 감정 분석값을 결합하여 두 모델 모두 같은 조건으로 주가 예측 모델을 개발하였다. 첫 번째와 두 번째 층에서 각각 50개의 유닛을 출력하도록 설정하였다. 또한 확률적 경사 하강법(SGD, Stochastic Gradient Descent)을 사용하고 학습률 0.01, 모멘텀 0.9으로 설정하였다. 학습률 0.01은 모델이 가중치를 업데이트할 때마다 이동할 단계의 크기를 의미한다. 모멘텀 0.9는 전 가중치 업데이트의 90%를 현재 업데이트에 반영함을 의미한다.

이 모델은 각각의 결합 값을 입력으로 받아, 5일 동안의 데이터를 학습하여 다음 2일 동안의 주가를 예측하는 방식으로 설정되었다. 본 연구에서는 epoch 값을 100,000, batch_size는 150, verbose는 0으로 설정하여 학습을 진행하였고, 이는 그림 7에서 확인할 수 있다.

date	Index	Stock-Sentiment Combined Index	X_{t-4}	X_{t-3}	X_{t-2}	X_{t-1}	X_t
2023-01-02	1467.03	879.9305					
2023-01-03	1452.47	871.1945					
2023-01-04	1421.19	852.4265					
2023-01-05	1443.72	865.9445					
2023-01-06	1411.34	846.5165					
2023-01-09	1459.05	875.1425					
2023-01-10	1497.83	898.4105					
2023-01-11	1501.47	900.5945					
2023-01-12	1535.87	921.2345					

그림 7. 예측 모델

Fig. 7. Prediction model

학습 횟수와 학습 기간을 조정해 가며 다양한 실험을 수행한 결과, 5일 동안의 데이터를 학습하여 다음 2일을 예측하고, 이를 총 100,000회 반복하는 설정이 가장 높은 예측 정확도를 보여주었다. 100,000회를 초과하여 학습을 진행할 경우, 예측 성능에 큰 변화가 없었거나, 일부에서 예측 성능이 오히려 저하되는 결과를 보였다. 이를 방지하기 위해 early stopping 기법을 활용하여 모델이 훈련 데이터에 과적합되지 않도록 가장 높은 예측 정확도를 보여주는 지점에서 학습을 멈추었다. 또한 과적합 문제를 완화하기 위해 정형 데이터와 비정형 데이터에 패널티 항을 추가하여 덜 중요한 정보의 영향을 받는 가중치를 축소하는 정규화 기법을 사용하였다[24]. 다음 표 1에서는 비중의 변화에 따른 결과값, 표 2에서는 학습 횟수에 따른 정확도 차이를 확인할 수 있다. 표 3에서는 학습 기간과 예측 기간에 따른 정확도 차이를 확인할 수 있다.

표 1. 비중별 결정계수

Table 1. Coefficient of determination by weight

stock weight	news weight	securities firm weight	LSTM (R^2)	GRU (R^2)
0.8	0.05	0.15	0.67641	0.82516
0.7	0.05	0.25	0.70161	0.79966
0.6	0.05	0.35	0.68615	0.84627

표 2. 학습 횟수별 결정계수

Table 2. Coefficient of determination by epoch

epoch	GRU(R^2)
1000	0.86412
5000	0.89817
10000	0.89471
100000	0.91659
200000	0.91054
300000	0.91382

표 3. 학습 횟수별 결정계수

Table 3. Coefficient of determination by time steps, periods

(learning days, prediction days)	GRU(R^2)
(5,2)	0.86412
(10,2)	0.85355
(10,5)	0.84449

4.3 모델 평가

감정분석을 활용한 주가지수 예측에 LSTM과 GRU 모델이 사용되었고, 평가 지표 MAE, RMSE, RMSLE, R^2 에 대해서 유사한 성능을 보였다. 표 4는 GRU 모델과 LSTM 모델의 결과값이다.

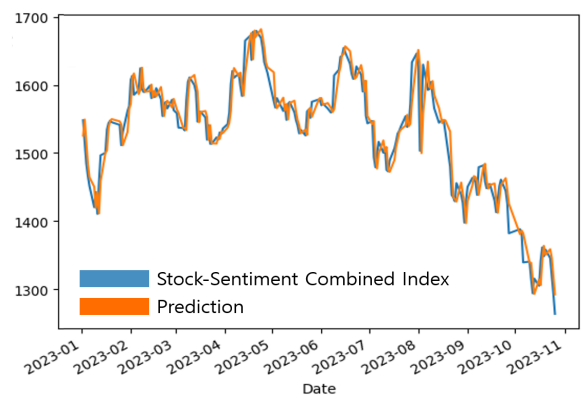
표 4. GRU 모델과 LSTM 모델의 결과값

Table 4. Result Values for GRU and LSTM

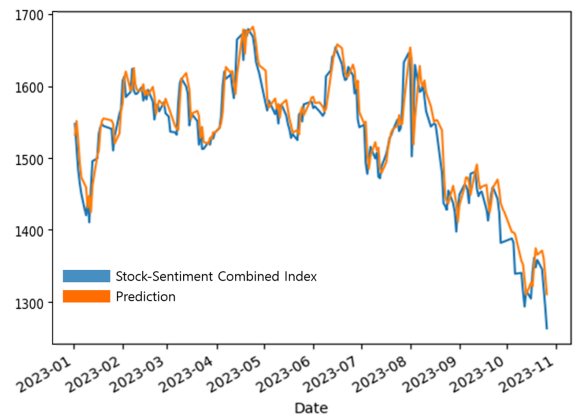
Model	MAE	RMSE	RMSLE	R^2	Time
GRU	11.163	14.790	0.0162	0.91835	74m
LSTM	11.780	15.486	0.0170	0.91049	122m

이 두 모델이 유사한 성능을 보인 이유는 두 모델 모두 시계열 데이터를 처리하는데 효과적인 RNN의 변형 모델이기 때문이다. 하지만 모든 평가 지표에서 GRU 모델이 LSTM 모델보다 주가지수 예측에 적합한 결과값이 도출되었다. 첫째, GRU 모델은 LSTM 모델보다 0.617 더 낮은 MAE를 가지며, 이는 GRU 모델의 예측이 평균적으로 실제 값에 더 가깝다는 것을 나타낸다. 둘째, GRU 모델은 LSTM 모델과 비교하였을 때 0.696 더 낮은 RMSE 값을 보여 GRU 모델이 전반적인 예측 정확도가 더 우수함을 나타낸다. 셋째, GRU 모델은 LSTM 모델과 비교하였을 때 더 높은 R^2 값이 도출되었다. 이

는 GRU 모델이 데이터의 더 많은 분산을 설명하고 데이터를 더 잘 적합함을 의미한다. 넷째, GRU 모델이 LSTM 모델보다 약 48분 더 짧은 시간 동안 학습을 완료하였다. 따라서, 비슷한 성능을 내면서도 학습 시간이 더 짧은 GRU 모델이 감정분석을 활용한 주가지수 예측에 더 우수하다고 할 수 있다. 하지만 이는 본 연구에서 사용한 특정 데이터셋과 설정에 국한된 결과일 수 있으며, 다른 조건이나 데이터셋에서는 그림 8과 같이 LSTM이 더 우수한 성능을 보일 수 있다.



(a) GRU



(b) LSTM

그림 8. 예측 결과값

Fig. 8. Prediction results

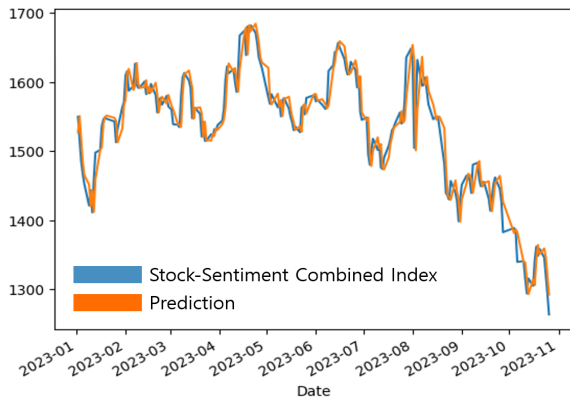
또한 주식 데이터를 다양한 시점에서 모델링하여 정형, 비정형, 주가 데이터의 비율을 조정하여 관련성이 낮은 정보의 영향력을 효과적으로 최소화하고 중요한 데이터의 영향력을 극대화하였다. 이러한 미세 조정 프로세스를 통해 기존 연구[25]에서 제기되었던 문제를 해결하고 예측 모델의 전반적인 성능과 정밀도를 향상시켰다.

LSTM과 GRU 모델을 주가 데이터에만 적용한 결과는 표 5와 그림 9에서 확인할 수 있다.

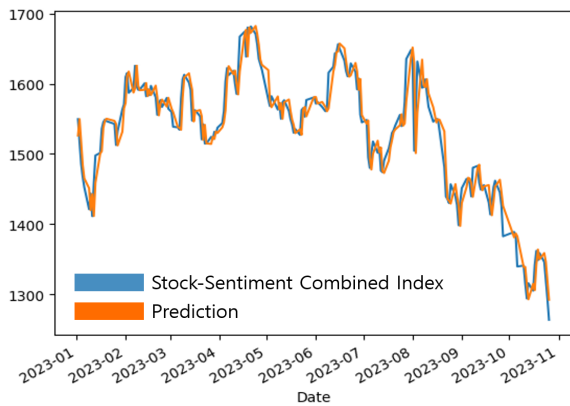
표 5. 주가만을 반영한 모델과 감정 분석값을 포함한 모델의 결과값

Table 5. Results from a model that includes only stock price and a model that includes sentiment analysis

Model		MAE	RMSE	RMSLE	R ²
stock + senti ment	GRU	11.163	14.790	0.0162	0.91835
	LSTM	11.780	15.486	0.0170	0.91049
stock	GRU	18.103	24.024	0.0157	0.92244
	LSTM	18.138	24.116	0.0157	0.92188



(a) GRU



(b) LSTM

그림 9. 주가 데이터만을 사용한 예측 결과값

Fig. 9. Prediction results using stock price data only

두 모델과 감정분석 값을 포함한 예측 모델을 비교했을 때, 주가만을 예측한 모델이 MAE와 RMSE 지표에서 더 높은 값을 나타냈다. MAE는 예측값과

실제값 사이의 평균 오차를 나타내며, RMSE는 큰 오차를 더 크게 반영하여 예측의 정확성을 평가한다. 특히 주가 예측에서는 예측 오류에 대한 민감도가 매우 중요한데, 이는 주가에 내재된 변동성을 고려한 것이다. 이는 감정 분석값을 포함한 예측 모델이 상황 변화에 더 민감하게 반응한다는 의미이다. LSTM과 GRU 모델을 주가 데이터에만 적용한 모델이 감정 기반 모델보다 높은 R²값(각각 0.92244 및 0.92188)을 보였으나, 이는 반드시 긍정적인 결과로 해석될 수만은 없다.

이러한 결과는 모델이 훈련 데이터에 과도하게 적합될 가능성을 내포하고 있다. 감정 분석값을 결합한 예측 모델이 주가 데이터만을 예측한 모델보다 R²값은 낮지만, MAE와 RMSE가 낮게 나왔다는 것은 이 모델이 새로운 데이터에 대해 더 잘 일반화되고 과적합의 위험을 줄인다는 것을 의미한다. 이는 감정 분석값을 포함한 주가 예측 모델이 실제 주가 예측에서 더 우수한 성능을 발휘함을 시사한다.

V. 결론 및 향후 과제

본 연구는 비정형 데이터와 정형 데이터를 활용하여 감정분석과 GRU 모델과 LSTM 모델 기반의 주가지수 예측 및 분석하였다, 이를 위해 2023년 1월부터 9월 30일까지의 기간 동안 자료를 수집하였으며, CLOVA Studio를 활용하여 데이터를 요약하였다. 또한, KR-FinBERT-SC를 이용하여 감정분석을 진행하였다. 이렇게 산출된 감정 점수와 주가지수를 결합하여 주가 예측 모델을 생성하였다.

본 연구의 결과는 주가지수 예측에 있어 감정분석의 활용 가능성을 확인하였다. 이는 감정분석이 주가 변동성을 예측하는데 필요한 중요한 요소로 작용할 수 있음을 보여주었다. 그러나 본 연구의 예측 모델은 주가지수의 비중, 비정형 데이터의 비중, 정형 데이터의 비중에 따라 그 성능이 크게 달라질 수 있음을 확인하였다. 이는 각 데이터의 중요도를 적절히 반영하는 것이 주가지수 예측의 정확도를 높이는 데 핵심적인 요소라는 점을 시사한다.

또한, 본 연구에서는 GRU 모델과 LSTM 모델의 성능 분석을 실시하여 GRU 모델이 주가지수에 대

한 예측 모델로서 좀 더 정확함을 알게 되었다.

LSTM과 GRU 모델의 비교에 관한 연구는 다수 있는데, 그 중 LSTM과 GRU를 활용한 프로그래밍 언어 예측 연구[26]에 따르면, 두 모델 모두 실제 데이터와 유사하게 보이지만 LSTM 같은 경우에는 급격한 변화에 대해 예측을 하지 못하는 반면, GRU는 급격한 변화에 반응한다고 보고하였다. 또한 데이터의 크기가 상대적으로 작은 경우, GRU 모델이 더 효율적이라는 점이 밝혀졌다. 김재호 등의 연구에 따르면 GRU 모델이 LSTM 모델에 비해 파라미터 수가 적기 때문에 복잡하지 않은 데이터에서 더 나은 성능을 발휘할 수 있다고 보고하였다[27]. 주요 성능 평가 지표인 MAE와 RMSE를 통해 GRU 모델이 LSTM 모델보다 낮은 오차를 보임으로써 높은 정확도를 나타내는 것으로 나왔다. 게다가 GRU 모델은 LSTM 모델에 비해 필요한 계산 시간이 짧은 것으로 나타나, 효율적인 측면에서도 우수함이 드러났다[28]. 이와 같은 연구들은 GRU 모델의 우수성을 입증하고 있는 연구인데, 본 연구에서도 마찬가지로 GRU 모델의 우수성을 입증하게 되었다.

하지만, 본 연구는 몇 가지 한계점을 가지고 있다. 첫째, 본 연구는 2023년의 한정된 기간의 데이터를 활용하였다. 따라서 다양한 시기의 데이터를 활용한 추가적인 연구가 필요하다. 둘째, 본 연구에서는 긍정과 부정 감정만을 고려하였다. 하지만 중립 감정 또한 주가에 영향을 미칠 수 있으므로 이를 고려한 추가적인 연구가 필요하다. 셋째, 본 연구에서는 LSTM 모델과 GRU 모델에만 값을 적용하였다. 이러한 모델 외에도 주가 예측에 더욱 효과적인 다른 모델들이 존재할 수 있다[29][30]. 넷째, 네이버 뉴스 기사로 제한한 한계점이 있다. 따라서 다양한 모델을 통한 주가 예측 성능의 비교 연구가 추가로 필요하다.

References

- [1] M. K. Kim, "The COVID-19 Crisis and the Recent Increase in Demand for Stock Investment", Korea Capital Market Institute, 2020.
- [2] M. K. Kim, "Individual Investors in the COVID-19 Phase : Investment Behavior and Investment Performance", Korea Capital Market Institute, 2021.
- [3] S. Hong, "Research on stock price prediction system based on BLSTM", Journal of the Korean Convergence Society, Vol. 11, No. 10, pp.19-24, Oct. 2020. <https://doi.org/10.15207/JKCS.2020.11.10.019>.
- [4] J. Heo and J. Yang, "Stock price prediction using SVM-based financial information", Journal of Information Science and Computing Practices, Vol. 21, No. 3, pp. 167-172, Mar. 2015. <https://doi.org/10.5626/KTCP.2015.21.3.167>.
- [5] J. Y. Park, J. Ryu, and H. J. Shin, "Predicting KOSPI stock index using machine learning algorithms with technical indicators", Journal of Information Technology and Architecture, Vol. 13, No. 2, pp. 331-340, Feb. 2016.
- [6] M.-S. Song and J. Min, "Comparison of stock price prediction using time series and non-time series data", Journal of the Korea Society of Computer and Information, Vol. 28, No.8, pp. 67-75, Aug. 2023. <http://dx.doi.org/10.9708/jksci.2023.28.08.067>.
- [7] Y. Kim, N. Kim, and S. R. Jeong, "Stock-index invest model using news big data opinion mining", Journal of Intelligence and Information Systems, Vol. 18, No. 2, pp.143-156, Feb. 2012. <https://doi.org/10.13088/JIIS.2012.18.2.143>
- [8] E.-J. Kim, S.-G. Chang, and S.-Y. T. Lee, "Big data news analysis in healthcare using topic modeling and time series regression analysis", Information Systems Review, Vol. 25, No. 3, pp.163-177, Mar. 2023. <http://dx.doi.org/10.14329/isr.2023.25.3.163>.
- [9] S.-W. Ahn and S.-B. Cho, "Stock prediction using news text mining and time series analysis", Korean Society of Information Scientists and Engineers Academic Presentation Papers, Vol. 37, No. 1C, pp. 364-369, Jun. 2010.
- [10] Y. I. Jang and Y. S. Jeong, "Prediction using a

- machine learning with news headlines for a pharmaceutical company during COVID-19 pandemic", *e-Business Studies*, Vol. 23, No. 4, pp. 237-255, Apr. 2022. <https://doi.org/10.20462/tebs.2022.8.23.4.237>.
- [11] E. A. Jang, H. R. Choi, and H. C. Lee, "Stock prediction using combination of BERT sentiment analysis and macro economy index", *Journal of the Korea Computer Information Society*, Vol. 25, No. 5, pp. 47-56, May 2020. <https://doi.org/10.9708/jksci.2020.25.05.047>.
- [12] J. W. Hyun, J. I. Lee, and H. K. Jo, "Sentiment analysis of news on corporation using KoBERT", *Accounting Research*, Vol. 47, No. 4, pp. 33-54, Apr. 2022. <https://doi.org/10.24056/KAR.2022.08.002>.
- [13] S. Lee, H. Jang, Y. Baik, S. Park, and H. Shin, "A small-scale korean-specific BERT language model", *Journal of KIISE*, Vol. 47, No. 7, pp. 682-692, Jul. 2020. <https://doi.org/10.5626/JOK.2020.47.7.682>.
- [14] KR-FinBert-SC, <https://huggingface.co/snunlp/KR-FinBert-SC> [accessed: Dec. 13, 2023]
- [15] CLOVA Studio, <https://ncloud.com/product/aiService/clovaStudio> [accessed: Dec. 13, 2023]
- [16] J. H. Jang and J. Y. Kim, "KOSPI prediction based on news sentiment analysis using KR-FinBert", *Proceeding of Symposium of the Korean Institute of Communications and Information Sciences*, pp. 1142-1143, 2023.
- [17] N. H. Choi, K. S. Kim, and H. S. Jang, "Prediction of initial price formation and impact analysis of public offering stock listing date using news and comment text: based on XGBoost", *The Korean Journal of Financial Management*, Vol. 40, No. 1, pp. 81-104, Jan. 2023. <https://doi.org/10.22510/kjofm.2023.40.1.005>.
- [18] T. Y. Lee, "The Effect of Investor Sentiment on the Real Estate Market and Stock Market : with a Focus on Unstructured Data and Vector Error Correction Model", Ph.D Dissertation, Graduate School of Kangnam University, 2020.
- [19] H. Kim, G. Joo, and H. Im, "Product category classification using word embedding and GRUs", *Journal of Korean Institute of Information Technology*, Vol. 19, No. 4, pp. 11-18, Apr. 2021. <http://dx.doi.org/10.14801/jkiit.2021.19.4.11>.
- [20] H. S. Jang, "Reduction of 3D information noise by using LSTM network", *Journal of Korean Institute of Information Technology*, Vol. 19, No. 8, pp. 91-98, Aug. 2021. <http://dx.doi.org/10.14801/jkiit.2021.19.8.91>.
- [21] D. Shin, G. Choi, and C. B. Kim, "Deep learning model for prediction rate improvement of stock price using RNN and LSTM", *Journal of Korean Institute of Information Technology*, Vol. 15, No. 10 pp. 9-16, Oct. 2017. <http://dx.doi.org/10.14801/jkiit.2017.15.10.9>.
- [22] C. Kwon, "A Complete Guide to Python Machine Learning", Revised 2nd edition. Wikibooks, 2022.
- [23] J. Hwang and C. Kim, "Deep learning parallel model to improve stock price prediction rate using technical analysis and environmental factors", *The Journal of Korean Institute of Information Technology*, Vol. 21, No. 11, pp. 53-61, Nov. 2023. <https://doi.org/10.14801/jkiit.2023.21.11.53>.
- [24] Y. Kawk and J. Huh, "Random augmentation technique for mitigating overfitting in neural networks for financial time series forecasting", *Journal of The Korean Data Analysis Society*, Vol. 25, No. 5, pp. 1653-1669, May 2023. <https://doi.org/10.37727/jkdas.2023.25.5.1653>.
- [25] H. Kwon, "A Comparative Study on the Effectiveness of Sentiment Score and Prediction Model in Stock Price Forecasting", Master's Thesis, Inje University, 2023.
- [26] M. Lee, D. Han, J. Bae, H. Seo, and W. Na, "Programming language prediction using LSTM and GRU", *Proceedings of the 2023 Korea Communications Society Winter Conference*, pp. 373-374, 2023.
- [27] J. Kim and Jang. Y. Kim, "Comparative analysis of the performance of BI-LSTM and GRU

algorithms for predicting the number of confirmed coronavirus cases", Journal of the Korea Information and Communication Society, Vol. 26, No. 2, pp. 187-192, Feb. 2022. <http://doi.org/10.6109/jkiice.2022.26.2.187>.

[28] S. Lee and K. Kwak, "Development of a demand prediction model for public bicycles in seoul using the GRU technique", Intelligent Information Research, Vol. 28, No. 4, pp. 1-25, Apr. 2022. <http://dx.doi.org/10.13088/jiis.2022.28.4.001>.

[29] H. Song and H. Choi, "Forecasting stock market indices using the recurrent neural network based hybrid models: Cnn-lstm, gru-cnn, and ensemble models", Applied Sciences, Vol. 13, No.7, Jul. 2023. <https://doi.org/10.3390/app13074644>.

[30] S. Narayana, S. N. D. Sri, S. R. Kumar, T. Ajay, and S. S. Vasiq, "Predicting the stock market index using GRU for the year 2020", 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), Bhubaneswar, India, Feb. 2024. <https://doi.org/10.1109/ESIC60604.2024.10481534>.

박 찬 정 (Chan Jung Park)



1988년 2월 : 서강대학교

전자계산학과(공학사)

1990년 2월 : KAIST 전산학과

(공학석사)

1998년 2월 : 서강대학교 대학원

전자계산학과(공학박사)

1990년 3월 ~ 1994년 2월 :

한국통신 소프트웨어연구소 전임연구원

1998년 2월 ~ 1999년 9월 : 한국통신 멀티미디어연구소
전임연구원

2011년 1월 ~ 2011년 12월 : UC Berkeley 방문학자

2013년 4월 ~ 2015년 2월 : 제주대학교 교육과학연구소
소장

1999년 9월 ~ 현재 : 제주대학교 컴퓨터교육과 교수

관심분야 : STEAM, 기술기반 교육, 에듀테크, 빅데이터
처리, 데이터마이닝

저자소개

양 혁 진 (Hyeok-Jin Yang)



2019년 3월 ~ 현재 : 제주대학교

컴퓨터교육과 학사과정

관심분야 : 정보교육, 인공지능,

빅데이터처리, 데이터마이닝

최 종 민 (Jong-Min Choi)



2020년 3월 ~ 현재 : 제주대학교

컴퓨터교육과 학사과정

관심분야 : 정보교육, 에듀테크,

인공지능, 빅데이터처리