

# 변이형 오토인코더와 어텐션 메커니즘을 결합한 차트기반 주가 예측

## Chart-based Stock Price Prediction by Combining Variation Autoencoder and Attention Mechanisms

배 상 현 (Sanghyun Bae) KCNET 데이터 사업부 대리  
최 병 구 (Byounggu Choi) 국민대학교 경영학부 교수, 교신저자

### 요 약

최근 인공지능 기법을 활용하여 캔들스틱 차트를 분석함으로써 주시가격 예측의 정확성을 높이고자 하는 다양한 연구가 진행되어 왔다. 그러나 이러한 연구들은 주시가격 예측을 위한 학습에 있어 캔들스틱 차트의 시계열적 특성을 고려하지 못한다는 점과 시장 참여자들의 감정 상태를 고려하지 못한다는 점 등이 문제로 지적되고 있다. 본 연구에서는 시장 참여자들의 감정상태를 반영하기 위해 변동성지수(VIX: volatility index) 차트를 캔들스틱 차트와 함께 고려하여 학습시키고 이를 변이형 오토인코더(VAE: variational auto encoder)와 어텐션 메커니즘(attention mechanisms)을 결합한 새로운 방법으로 분석하여 캔들스틱 차트의 시계열적 특성을 고려함으로써 기존 연구의 한계를 극복하고자 한다. 본 연구에서 제안한 방법의 성능 비교를 위해 S&P 500 기업 가운데 50개를 임의로 추출하여 제안한 방법을 통해 이들의 주시가격을 예측하고 이를 합성곱 신경망(CNN: convolutional neural network) 또는 장단기메모리(LSTM: long-short term memory) 등과 같은 기존 방법들과 비교하였다. 비교 결과 기존 방법들에 비해 본 연구에서 제안한 방법이 더 우수한 성능을 보이는 것으로 나타났다. 본 연구는 시장 참여자들의 감정 상태와 캔들스틱 차트의 시계열적 특성을 고려함으로써 주시 가격 예측의 정확성을 높였다는 점에서 그 의의가 있다.

**키워드 :** 캔들스틱 차트, 변이형 오토인코더, 어텐션 메커니즘, 변동성지수 차트, 주시가격 예측

## I. 서 론

주시시장은 다양한 변수에 의해 영향을 받을 뿐 아니라 많은 잡음이 포함되어 있어 이를 예측하는 것은 매우 어려운 과제이다. 더욱이 주시시장의 비정상성(non-stationarity)과 비선형성(non-linearity)으로 인해 이의 변동이나 수익을 예측하

는 것은 오랜 기간 해결되지 않은 경영학의 주요 연구주제로 인식되어 왔다(신동하 등, 2017). 그럼에도 불구하고 미래의 주시가격을 예측함으로써 얻을 수 있는 이익의 크기가 너무도 크기 때문에 이를 예측하기 위한 끊임없는 노력이 이루어져 온 것도 사실이다(김현모, 박재홍, 2016).

주시가격 예측을 위한 분석 방법은 크게 기본적인

분석(fundamental analysis)과 기술적 분석(technical analysis)으로 구분할 수 있다. 기본적인 분석은 현재 경제 상태, 전체 산업 상황, 기업의 재무상태 등과 같이 주시가격에 영향을 미치는 기본 조건 및 결정 요인을 분석하여 미래 주시가격의 추세를 예측하는 방법인 반면 기술적 분석은 오직 주시가격 변동 가격 움직임 자체만을 분석하여 미래의 주시가격 추세를 예측하는 방법이다(Murphy, 2011). 기본적인 분석이 주시가격 변동의 원인을 분석하는 방법이라면 기술적 분석은 원인보다는 현재 나타난 결과에 집중하는 분석이라 할 수 있다.

Fama(1965)가 제안한 효율적 시장 가설(efficient market theory)에 따르면 주시가격에 영향을 미치는 모든 정보는 이미 현재 주시가격에 반영되어 있기 때문에 기본적인 분석이나 기술적 분석 모두 미래 주시가격 예측에는 크게 도움이 되지 않는다. 그러나 점점 더 많은 연구들은 주시 시장을 효율적 시장이 아닌 비효율적 시장으로 간주하고 기술적 분석 혹은 기본적인 분석을 통해 주시가격을 예측하고자 하고 있다. 예를 들면 Marshall *et al.*(2006)은 과거 주시가격 데이터를 기반으로 미래 주시가격을 예측함으로써 높은 수익률을 얻을 수 있음을 실증하였다. 이외에도 다양한 연구자들이 기술적 혹은 기본적인 분석을 활용함으로써 수익률 향상을 가져올 수 있음을 주장하고 있다(Braun and Chandler, 1987; Ou and Penman, 1989).

빅데이터와 인공지능 기술이 발달함에 따라 연구자들은 주시가격 예측에 이러한 기술을 접목함으로써 주시가격의 예측력을 높이고자 다양한 시도를 하고 있다(Patalay and Bandlamudi, 2020). 특히 최근에는 예측의 적시성과 기업 재무제표의 신뢰성 문제가 지적되고 있는 기본적인 분석보다는 주시가격의 변동 추세를 기반으로 미래 주시가격을 예측하는 기술적 분석이 각광받고 있다. 기술적 분석은 주시가격의 변동성 예측을 위해 다양한 차트를 활용하는데 특히 캔들스틱(candlestick) 차트가 많이 활용되고 있다. 캔들스틱 차트는 수치 데이터인 재무 데이터를 투자자들에게 의미 있는 패

턴으로 전달함으로써 패턴 인식을 선호하는 투자자에게 강력한 예측 방법을 제공할 뿐 아니라 추세 변화의 시점을 빠르게 파악할 수 있다는 장점도 제공하고 있다. 이에 따라 많은 연구들이 인공지능 기법을 활용하여 캔들스틱 차트를 분석함으로써 주시가격을 예측하고자 시도하였다. 예를 들면, Guo *et al.*(2018)은 합성곱 신경망(CNN: convolutional neural network)을 사용하여 캔들스틱 차트를 분석함으로써 주시가격의 등락을 예측하고자 하였으며, Zhipeng and Chao(2019)는 서포트 벡터 머신(SVM: support vector machine)을 활용하여 캔들스틱 차트를 분석함으로써 주시가격의 등락 예측을 시도하였다.

그러나 캔들스틱 차트를 활용한 기존 인공지능 연구들은 크게 i) 주시가격 예측을 위한 학습 과정에서의 과적합(overfitting) 위험이 높다는 점, ii) 캔들스틱 차트의 시계열적 특성을 고려하지 못한 학습이 이루어지고 있다는 점, iii) 시장 참여자들의 감정 상태를 고려하지 못한다는 점 등이 단점으로 지적되고 있다. 이를 보다 구체적으로 살펴보면 첫째, 기존 연구들은 주시가격 예측을 위한 특성 추출(feature extraction)과 차원 축소(dimension reduction)를 위해 오토인코더(AE: autoencoder) 기법을 주로 활용하였다. 그러나 오토인코더 기법은 주시가격 예측에 필요하지 않은 특성까지 학습함으로써 과적합의 위험이 상대적으로 높은 것도 사실이다(Chu and Cai, 2017). 따라서 정확한 주시가격 예측을 위해서는 이러한 오토인코더 기법의 한계점 극복이 필요하다. 둘째, 기존 연구들은 시계열적 특성을 가진 주시가격의 변화를 적절하게 고려하지 못하고 있다. 주시가격의 변화와 같이 고려해야 하는 시계열 기간이 길 경우 장기 의존성(long term dependencies) 문제가 발생한다. 이러한 장기 의존성 문제 해결을 위해 장단기 메모리(LSTM: long-short term memory) 방법이 개발되어 활용되고 있으나 이 방법 역시 장기 의존성 문제에서 완전히 자유롭지 못하다는 지적을 받고 있다(Bao *et al.*, 2017). 장기 의존성은 주시가격 예측을 위한 학

습의 성능을 저해시키고 결과적으로 예측의 정확성을 저해한다. 따라서 예측의 정확성을 향상시키기 위해서는 이러한 한계점을 극복할 필요가 있다. 마지막으로 기존 연구들은 시장 참여자들의 감정 상태가 주시가격에 미치는 영향이 매우 중요함에도 불구하고 이를 고려하지 못하고 있다(Ding *et al.*, 2015.; Huynh *et al.*, 2017; Nguyen *et al.*, 2015). 행동경제학 이론에 따르면 대중의 감정은 개인의 행동과 의사결정에 큰 영향을 미칠 수 있다. 이를 주식 시장에 적용하면 시장 참여자들의 감정은 개별 투자자들의 행동에 영향을 미쳐 궁극적으로 개별 주시가격에 영향을 미친다고 할 수 있다. 따라서 주시가격의 정교한 예측을 위해서는 시장 참여자들의 감정을 고려하는 것이 중요하다.

본 연구에서는 기존 연구의 이러한 한계점을 보완하기 위해 i) 어텐션 매커니즘(attention mechanism)과 변이형 오토인코더(VAE: variational autoencoder)를 결합한 새로운 분석 기법을 제안하고 ii) 시장 참여자들의 감정 상태를 고려하기 위해 변동성지수(VIX: volatility index) 차트와 캔들스틱 차트를 함께 고려함으로써 주시 가격의 예측 성과를 높이고자 한다. 본 연구는 기존 연구와 달리 변이형 오토인코더를 활용함으로써 오토인코더의 과적합 위험을 해결할 수 있다는 점, 어텐션 메커니즘 기법을 양방향(bi-directional) 장단기메모리에 적용함으로써 장단기메모리의 장기 의존성 문제점을 해결할 수 있다는 점, 변동성지수 차트의 고려를 통해 기존 캔들스틱 차트 연구들이 고려하지 못했던 시장 참여자들의 감정 상태를 고려할 수 있다는 점에서 기존 연구와 차별 점이 있다.

본 연구는 다음과 같이 구성된다. 다음 장에서는 인공지능을 활용한 기존 주시가격 예측 연구들을 요약하고 관련 인공지능 기법을 소개한다. 제 III장에서는 연구 모형, 데이터 셋(data set), 평가 방법 등을 포함한 연구방법을 설명한다. 제 IV장에서는 데이터를 분석하고 분석결과와 이의 함의를 논의한다. 마지막 장에서는 본 연구의 결론 및 향후 연구과제를 제안한다.

## II. 선행 연구

### 2.1 인공지능을 활용한 기존 주시가격 예측 연구

최근 빅데이터 및 인공지능 관련 기술이 비약적으로 발전함에 따라 다양한 인공지능-특히 딥러닝(deep learning)-기법에 대한 연구가 활발하게 진행되고 있다. 주시가격 예측 연구도 이러한 추세에 부응하여 다양한 인공지능 기법을 활용함으로써 주시가격 예측의 정확도를 높이고자 하는 시도가 이루어지고 있다.

인공지능을 활용한 주시가격 예측 연구는 크게 i) 정형적인 주시가격 데이터를 활용한 연구, ii) 비정형적인 시장의 감성 데이터를 활용한 연구, iii) 비정형적인 이미지 데이터를 활용한 연구로 구분할 수 있다. 주시가격 예측을 위해 정형적인 주시가격 데이터를 활용한 초기 인공지능 연구는 컴퓨팅 파워 등의 한계로 인해 단순 주식시장 변수들을 입력 값으로 하는 인공 신경망(ANN: artificial neural network)을 주로 활용하였다(Niaki and Hoseinzade, 2013). 그러나 기술의 발달과 더불어 합성곱 신경망(CNN)을 사용하거나 순환 신경망(RNN: recurrent neural network)을 사용하여 주시가격 예측을 시도하는 연구들이 발전하기 시작하였고, 과거 딥러닝 기반 주시가격 예측 연구에 사용되었던 모델에 전통적인 특성 엔지니어링(feature engineering) 방식을 혼합하는 방법 또한 자주 활용되고 있다(Persio and Honchar, 2016; Song, 2018; Zhong and Enke, 2017). 또한 최근엔 합성곱 신경망 기법과 유전 알고리즘(GA: genetic algorithm)을 결합한 연구(Chung and Shin, 2020)와 시가, 종가, 최고가, 최저가의 관계가 밀접하다는 것을 이용하여 여러 장단기메모리를 결합하여 주가를 예측한 연구(Ding and Qin, 2020) 등이 이루어지고 있다.

비정형적인 시장의 감성 데이터를 활용한 연구는 자연어 처리 기술의 발전과 더불어 나타나기 시작하였다(Nguyen *et al.*, 2015; Pagolu *et al.*, 2016).

자연어 처리 분석 기법이 발달함에 따라 사람들이 쓴 글로부터 감정 상태를 어느 정도 예측할 수 있게 되었다. 이에 따라 시장 참여자들이 주식 시장과 관련하여 작성한 글로부터 감정 상태를 추출하고 이를 분석하여 주시가격 예측을 시도하는 연구들이 이루어지고 있다. Ding *et al.*(2015)은 경제 신문의 내용에 있는 사건들과 해당 기업을 하나의 튜플(tuple)로 묶어 이벤트 임베딩(event embedding)으로 벡터화하고 딥러닝 모델을 활용하여 이를 분석함으로써 다음 날 주시가격의 등락을 예측하고자 하였다. 뉴스 뿐 아니라 소셜네트워크서비스에 올라오는 글들의 감정 점수를 이용하여 주시가격 예측을 시도한 연구도 다수 이루어져 왔다. 예를 들면, Nguyen *et al.*(2015)은 잠재 디리클레 할당(LDA: latent dirichlet allocation) 방법을 활용하여 트위터 게시물을 분석함으로써 보다 개선된 주시가격 예측이 가능함을 실증하였다. Huynh *et al.*(2017)은 시장 참여자들의 감정 상태와 주시 시장의 관계를 시계열 형태로 분석함으로써 예측의 성능을 개선하였다.

최근 이미지 분석 기술이 발달함에 따라 주시가격의 변동을 이미지로 나타내는 차트를 기반으로 주시가격을 예측하고자 하는 연구가 각광 받기 시작하였다. Bao *et al.*(2017)은 캔들스틱 차트와 다양한 기술적 지표를 웨이블릿 변환(wavelet transform), 오토인코더, 장단기메모리의 입력 값으로 활용하여 주시가격을 예측하고자 하였다. Kim and Kim (2019)은 캔들스틱 차트와 캔들스틱 차트의 구성 데이터를 합성곱 신경망과 장단기메모리의 입력 값으로 활용하여 주시가격을 예측하고자 하였다. 나아가 Liu and Song(2018)은 캔들스틱 차트를 다양한 딥러닝 모델로 분석함으로써 캔들스틱 차트와 딥러닝 모형의 적합성(fit)을 파악하고자 하였다.

인공지능을 활용한 주시가격 예측 관련 기존 연구는 정형적인 주시가격 데이터를 기반으로 인공 신경망(Niaki and Hoseinzade, 2013), 합성곱 신경망(Chung and Shin, 2020), 순환 신경망(Song, 2018), 장단기메모리(Ding and Qin, 2020) 등 다양

한 기법들을 활용해 주시가격 예측의 정확성을 높이고자 시도하였다. 이러한 정형적인 주시가격 데이터 이외에 비정형적인 시장의 감정 데이터를 활용한 연구는 경제 신문(Ding *et al.*, 2015), 소셜네트워크서비스(Nguyen *et al.*, 2015) 등의 활용을 통해 합성곱 신경망, 잠재 디리클레 할당 등의 모형을 활용하여 분석함으로써 기존 정형적인 주시가격 데이터 활용 연구의 단점을 보완하고자 하였다. 나아가 최근에는 이미지 분석 기술의 발달에 따라 캔들스틱 차트 등의 이미지를 합성곱 신경망, 장단기메모리 등과 같은 다양한 딥러닝 모델로 분석하여 주시가격 예측의 정확성을 높이는데 일정 정도 기여하고 있다(Bao *et al.*, 2017; Kim and Kim, 2019). <표 1>은 딥러닝 기반 주시가격 예측 관련 기존 연구를 요약한 표이다.

기존 연구를 통합 분석해 보면 다음과 같은 흥미로운 사실을 발견할 수 있다. 첫째, 대다수의 연구들이 시장 참여자들의 감정 상태를 고려하지 않고 있다(Bao *et al.*, 2017; Kim and Kim, 2019; Liu and Song, 2018; Niaki and Hoseinzade, 2013; Persio and Honchar, 2016). 시장 투자자들의 감정 상태가 주시가격에 영향을 미침에도 불구하고 이를 고려하지 않음으로써 미래 주시가격을 정확하게 예측하지 못하는 한계를 보이고 있다. 둘째, 감성을 고려한 연구들의 경우 뉴스나 소셜네트워크서비스의 게시글에 중점을 두고 있다(Ding *et al.*, 2015; Huynh *et al.*, 2017; Nguyen *et al.*, 2015). 이러한 뉴스나 소셜네트워크서비스 게시글의 경우 일정 정도 시장 참여자의 감성을 나타내고 있지만 시장 참여자들의 진정한 감성을 파악하기 어렵다는 점에서 문제가 있다. 셋째, 대다수의 연구들이 활용하고 있는 인공지능 분석 기법이 주시가격 학습 과정에서의 과적합이나 주시가격의 시계열적 특성을 명확하게 고려하지 못하고 있다는 방법론적인 한계점을 보이고 있다(Guo *et al.*, 2018; Song, 2018; Zhong and Enke, 2017). 이로 인해 미래 주시가격 예측의 정확성이 실무자들의 기대에 미치지 못하고 있는 실정이다.

〈표 1〉 인공지능 활용 주시가격 예측 기존 연구 요약

구분	저자	데이터셋		분석모델	예측형태	분석기간	
정형	기본	Niaki and Hoseinzade(2013)	S&P 500	Technical indexes	ANN	가격상승 또는 하락	1994.3 ~ 2008.6
		Persio and Honchar (2016)	S&P 500	Technical indexes	ANN, WCNN, RNN	가격상승 또는 하락	2006 ~ 2016
		Zhong and Enke (2017)	S&P 500	Technical indexes	ANN, PCA	가격상승 또는 하락	2003.6 ~ 2015.5
		Song(2018)	S&P 500	Technical indexes	RNN, SVM,XGBoost	가격상승 또는 하락	2010.1 ~ 2017.12
		Chung and Shin (2020)	KOSPI index		GA, CNN	가격자체	2000.1 ~ 2016.12.31
		Ding and Qin (2020)	Shanghai composite index Petro China ZTE		LSTM	가격자체	NA
비정형	감성	Ding <i>et al.</i> (2015)	S&P 500	Finance news	EB, CNN	가격상승 또는 하락	2006.10 ~ 2013.11
		Nguyen <i>et al.</i> (2015)	Twitter, price(OHLC)		LDA	가격상승 또는 하락	2012.7 ~ 2013.7
		Huynh <i>et al.</i> (2017)	S&P 500	Reuters News dataset	GRU	가격상승 또는 하락	2006.10 ~ 2013.12
	이미지	Bao <i>et al.</i> (2017)	S&P 500; Hang Seng index; CSI 300; Nikkei 225; Nifty 50	Candlestick chart, Technical indexes	AE+LSTM	가격자체	2008.7 ~ 2016.9
		Guo <i>et al.</i> (2018)	TAIFEX	Candlestick chart	AE+CNN	가격상승 또는 하락	1998 ~ 2016
		Liu and Song(2018)	10 China Market index stock; 10 White Horse stock	Candlestick chart	SVM, CNN, DNN	가격상승 또는 하락	2006.1 ~ 2017.8
		Kim and Kim(2019)	S&P 500	Candlestick chart	CNN, LSTM	가격자체	2016.10 ~ 2017.10
	감성 + 이미지	본 연구	S&P 500	Candlestick chart, VIX chart	VAE+ Attention model and Dense layer	가격상승 또는 하락	1993.7 ~ 2019.7

ANN: artificial neural network, WCNN: wavelet CNN, PCA: principal component analysis, XGBoost: extreme gradient boosting, EB: event embedding, GRU: gated recurrent unit, DNN: deep neural network

## 2.2 변이형 오토인코더

Kingma and Welling(2013)이 제안한 변이형 오토인코더는 비지도 기반 생성 모델로 오토인코더와 베이저안(Bayesian) 추론을 바탕으로 데이터의 분포를 추정하는 방법이다. 오토인코더의 잠재 변수는 인코더(encoder)에 의해 압축된 입력 데이터의 특징이라고 할 수 있는 반면 변이형 오토인코더의 잠재 변수는 특정한 분포를 따르는 입력 데이터의 특징이라고 할 수 있다. 변이형 오토인코더의 대략적인 모형은 <그림 1>과 같다.

오토인코더의 손실함수(loss function)는 입력 데이터(Input)  $X$ 와 잠재 변수  $z$ 가 디코더(decoder)를 거쳐 나온 출력 데이터(Output)간의 차이로 파악하며 이는 다음의 식 (1)과 같이 나타낼 수 있다. 오토인코더는 이 손실함수를 최소화하도록 학습하게 되며 잠재 변수  $z$ 는 입력 데이터  $X$ 의 특징을 보존하고 있으므로 입력 값의 유사도를 구하거나 분류 문제에 사용되게 된다.

$$\begin{aligned} & \text{binary cross entropy}(X, X') \quad (1) \\ &= -\frac{1}{N} \sum_i (X_i \log(X'_i) + (1 - X_i) \log(1 - X'_i)) \end{aligned}$$

이와 달리 변이형 오토인코더는 잠재 변수  $z$ 가 확률 분포를 따르게 하기 위해 쿨백-라이블러 발

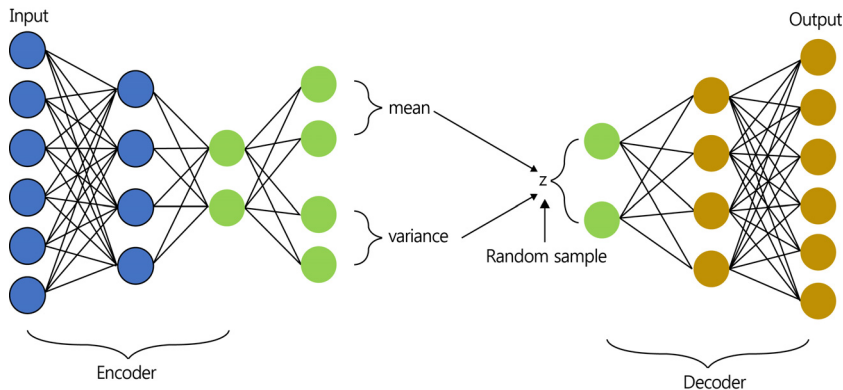
산(Kullback-Leibler divergence)을 사용한다. 쿨백-라이블러 발산이란 두 확률 분포의 차이를 계산하는 데 사용하는 함수로 다음 식 (2)와 같이 나타낼 수 있으며, 이는 어떤 이상적인 분포에 대해 그 분포를 근사하는 다른 분포를 사용해 샘플링을 했을 때 발생할 수 있는 정보 엔트로피 차이를 의미한다.

$$\begin{aligned} & KL[\mathcal{N}(\mu, \sigma^2) \parallel \mathcal{N}(0, 1)] \quad (2) \\ &= -\frac{1}{2}(1 + \log \sigma - \mu^2 - \sigma^2) \end{aligned}$$

변이형 오토인코더는 손실함수에 잠재 변수  $z$ 가 정규분포를 따르도록 다음과 같은 쿨백-라이블러 발산 값을 추가한다. 따라서 변이형 오토인코더의 손실함수는 다음의 식 (3)과 같은 형태로 나타난다.

$$\begin{aligned} & \text{Total Loss} = \text{binary cross entropy}(X, X') \quad (3) \\ & \quad + KL[\mathcal{N}(\mu, \sigma^2) \parallel \mathcal{N}(0, 1)] \end{aligned}$$

기존 캔들스틱 차트를 이용한 주식가격 예측 연구들은 특성 추출과 차원축소를 위해 오토인코더를 사용하였다. 그러나 앞서 언급한 바와 같이 오토인코더의 목적은 재구축 비용을 최소화하는 것이기 때문에 재구축에 필요하지 않은 특징까지



<그림 1> 변이형 오토인코더(Variational Autoencoder)

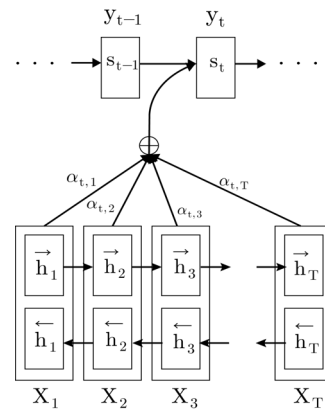
학습하는 과적합의 위험이 존재한다. 이와 달리 변이형 오토인코더는 새로운 데이터 생성을 위한 목적으로 잠재 변수  $z$ 가 확률 분포를 따르게 하는데, 이는 표준화(standardization)의 효과를 내어 오토인코와 비교하여 더 우수하게 입력 데이터  $X$ 의 특성을 추출할 수 있게 해 준다. 이에 따라 본 연구에서는 변이형 오토인코더를 사용함으로써 기존 오토인코더 방식의 한계점을 극복하고자 한다.

### 2.3 어텐션 메커니즘

인공지능 분석 방법 가운데 하나인 순환 신경망은 데이터를 순차적으로 입력함으로써 순서를 고려할 필요가 있는 시계열 데이터 분석에 널리 활용되었다(Karevan and Suykens, 2020; Liu et al., 2017). 그러나 순환신경망은 시계열 기간이 너무 길면 비용 함수의 편미분 계수가 0에 가까워 지거나 매우 커지는 문제가 발생하여 학습이 제대로 이루어지지 않게 된다(Bengio et al., 1993). 이러한 문제점을 해결하기 위해 Hochreiter and Schmidhuber(1997)는 장단기메모리 모델을 제안하였다. 장단기메모리의 경우 순환 신경망과 비슷한 구조를 가지지만 은닉 계층을 메모리 셀이라는 구조로 만듦으로써 순환 신경망의 한계를 극복하였다. 또한 순환 신경망은 다음 시점의 데이터가 있음에도 불구하고 이전까지의 데이터만을 가지고 예측을 함으로써 예측의 정확성이 상대적으로 약화되는 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해 과거에서 현재로의 데이터 흐름만을 고려하는 기존 순환 신경망을 개선하여 미래에서 현재로의 흐름도 함께 고려하는 양방향 순환 신경망이 제안되었다(Schuster and Paliwal, 1997). 나아가 Graves and Schmidhuber(2005)는 양방향 순환 신경망과 장단기메모리를 결합하여 양방향 장단기메모리를 제안하였다.

양방향 장단기메모리의 경우 순환 신경망이 가지고 있는 장기 의존성 문제와 다음 시점의 데이터를 활용하는 문제를 효과적으로 해결하여 준다.

그러나 이러한 양방향 장단기메모리는 예측에 있어 기존에 입력된 데이터 전체를 동일한 비율로 참고한다는 측면에서 여전히 문제점이 있다. 즉, 예측에 있어 입력 데이터의 가중치를 고려하지 못한다는 단점이 있다. 이러한 문제를 해결하기 위해 Bahdanau et al.(2015)는 어텐션 메커니즘을 제안하였다. 어텐션 메커니즘은 출력을 예측하는 매 시점마다 기존에 들어온 입력 데이터 전체를 다시 한번 참고하는데 전체를 동일한 비율로 참고하는 것이 아니라 모델로 하여금 예측에 있어 보다 중요한 부분을 참고하게 함으로써 예측의 정확성을 높일 수 있다. 따라서 양방향 장단기메모리에 어텐션 메커니즘을 적용함으로써 본 연구의 목적을 보다 효과적으로 달성할 수 있다. 본 연구에서 사용하는 어텐션 메커니즘의 대략적인 모형은 <그림 2>와 같으며 이를 보다 구체적으로 설명하면 다음과 같다.



<그림 2> 어텐션 메커니즘(Attention Mechanism)

$t$ 시점의 어텐션 층 은닉 노드값( $s_t$ )을 구하기 위해선  $t-1$ 시점의 은닉 값( $s_{t-1}$ )과 이전의 출력 값( $y_{t-1}$ )이 필요하며  $t$ 시점의 출력을 예측하기 위한 어텐션 값( $a_t$ )이 필요하다. 이를 식으로 나타내면 식 (4)와 같다.

$$s_t = f(s_{t-1}, y_{t-1}, a_t) \tag{4}$$

어텐션 값을 도출하기 위해서는  $t$ 시점의 출력 값( $y_t$ )이 필요하다. 이를 구하기 위해서는 입력 값의 은닉 노드( $h$ ) 각각이 어텐션 레이어의  $t-1$ 시점 은닉 값과 얼마나 유사 한지 판단하는 어텐션 스코어(attention score)를 구할 필요가 있다. 어텐션 스코어 값을 구하기 위해서는 어텐션 레이어의  $t-1$ 시점 은닉 값을 전치(transpose)하여 입력 값의 은닉 노드 각각과 내적(dot product)을 수행해야 한다(식 (5) 참조). 이렇게 얻은 어텐션 스코어 값(식 (6) 참조)에 소프트맥스(softmax) 함수를 적용하여 어텐션 분포(attention distribution)를 구할 수 있다(식 (7) 참조).

$$score(s_{t-1}^T, h_i) = s_{t-1}^T h_i \quad (5)$$

$$e^t = [s_{t-1}^T h_1, \dots, s_{t-1}^T h_N] \quad (6)$$

$$\alpha^t = softmax(e^t) \quad (7)$$

이렇게 얻어진 어텐션 분포( $a_t$ )를 입력 값의 은닉 노드와 곱하여 모두 더한 값이 어텐션 값이 된다(식 (8) 참조). 이렇게 구한 어텐션 값을 어텐션 레이어의  $t-1$ 시점 은닉 값과 결합(concatenate)함으로써 은닉 노드값( $s_t$ )을 구할 수 있게 된다(식 (9) 참조).

$$a_t = \sum_{i=1}^N \alpha^t h_i \quad (8)$$

$$s_t = f(v_t, y_{t-1}) \quad (9)$$

많은 연구에서 어텐션 메커니즘은 다른 인공지능 분석 방법에 비해 더 나은 성능을 보였으며 이로 인해 현재 이미지 분석을 포함한 다양한 분야에서 적극적으로 활용되고 있다(Rush *et al.*, 2015; Yao *et al.*, 2015).

### III. 연구 방법

#### 3.1 연구 모형

본 연구의 주요 목적을 달성하기 위해 먼저 개별 주식의 캔들스틱 차트를 변이형 오토 인코더에 입력 값으로 입력한 후 재구축 손실이 적어지도록

학습을 시킨다. 이렇게 학습된 변이형 오토 인코더를 이용하여 캔들스틱과 변동성지수 차트의 잠재(latent) 값을 구하게 된다. 이 잠재 값들을 어텐션 모델의 입력 값으로 활용하게 되면 변이형 오토인코더와 어텐션 모델을 활용하여 주식의 변동성을 예측하는 모델이 된다. 그러나 본 연구에서는 예측의 정확도를 보다 높이기 위해 상대적 중요도가 높을 것으로 판단되는 예측 전 20일치의 자료에 좀 더 가중치를 높이기 위해 밀집층(dense layer)을 추가하여 어텐션 모델과 밀집층을 이용한 앙상블(ensemble) 모형을 구축하였다. 다음으로 밀집층과 어텐션 모델에 잠재 값들을 각각의 입력 값으로 활용하여 개별 신경망의 출력물들을 결합(concatenate)한 후 이를 입력 값으로 사용하였다. 마지막으로 소프트맥스(softmax) 함수를 통과한 익일 증가(현재가)가 금일 증가보다 상승할 것인지 아니면 하락할 것인지를 예측함으로써 본 연구의 목적을 달성할 수 있게 된다. 이러한 연구 방법 및 절차를 요약하면 다음 <그림 3>과 같으며 본 연구에서는 이를 VATD(variational autoencoder with attention technic and dense layer)로 명명한다.

##### 3.1.1 단계 1: 데이터 셋 구축

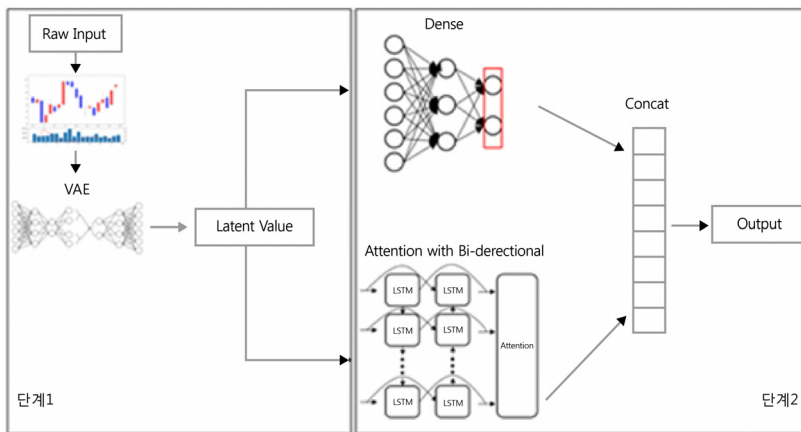
단계 1에서는 대상이 되는 차트의 특징을 변이형 오토인코더로 추출하고 추출된 특징을 바탕으로 단계 2에 입력이 되는 데이터 셋을 구축하게 된다. 데이터 셋을 구축하기 위해 캔들스틱 차트와 시장 참여자의 감성을 나타내는 변동성지수 차트를 하나의 이미지로 구성한다. 입력에 사용될 각각의 차트를 하나의 이미지로 만드는 이유는 개별적으로 입력된 차트에 비해 하나의 이미지에 다수의 차트를 넣는 것이 성능 향상에 보다 효과적이기 때문이다(Kim and Kim, 2019). <그림 4>는 본 연구에서 입력으로 활용되는 데이터 셋의 예시이다.

이렇게 만들어진 차트는 변이형 오토인코더의 입력 데이터가 되고 변이형 오토 인코더를 활용하여 단계 2에서 사용될 잠재 값들을 추출하게 된다. 변이형 오토인코더에 입력되는 이미지의 크기는 넓이와

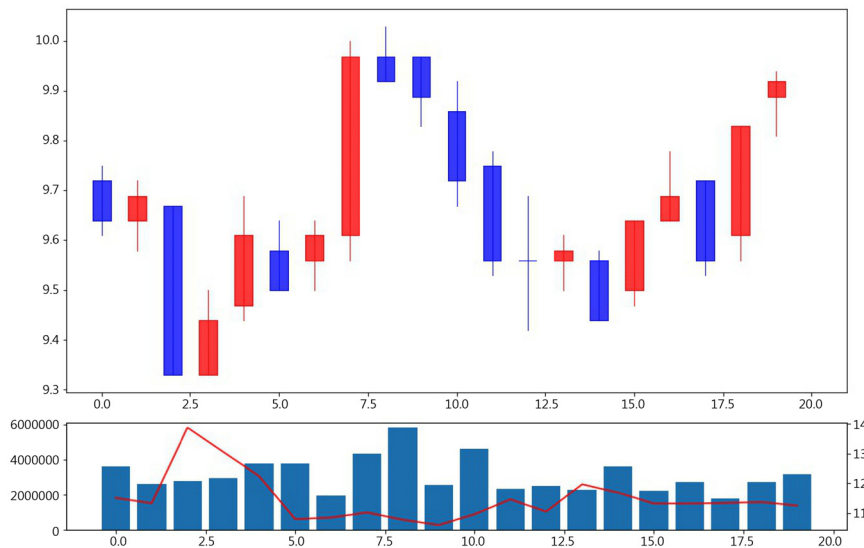


높이 모두 128로 설정하였고 변이형 오토인코더의 네트워크 구조는 다음과 같이 구성하였다. 첫째, 인코더 부분에 다섯 층의 합성곱(convolution) 네트워크를 쌓고 각 합성곱 층의 필터 수는 입력층에 가까운 순서대로 128, 256, 256, 512, 1,024개로 설정하였다. 필터의 크기는 3, 스트라이드(stride)는 모두 2로 설정하였다. 합성곱 층을 거쳐 나온 결과에 평탄화 층을 적용한 후 밀집층을 연결하였고 밀집층 노드 수와

변이형 오토인코더의 잠재 차원(latent dimension)은 각각 30으로 설정하였다. 둘째, 디코더 층은 인코더 층의 역순으로 전치 합성곱(transpose convolution) 층을 쌓았으며 이에 따라 각 필터의 개수는 1,024, 512, 256, 256, 128개가 된다. 사용할 모든 층의 활성화 함수는 렐루(ReLU: rectified linear unit) 함수를 사용하였고 최적화 알고리즘은 아담(Adam: adaptive moment estimation) 방식을 사용하였다.



〈그림 3〉 연구 절차



〈그림 4〉 입력 차트 예시

<표 2> 구성된 입력 데이터 셋 예시

	0	1	...	329
0	-106.03342	2.5770082	...	1.3704153
1	11.75793	-0.1898792	...	0.9307256
2	-0.7532625	0.81035775	...	0.9168761
...	...	...	...	...

이러한 과정을 거쳐 나온 잠재 값들은 밀집층에 입력되는 데이터 셋(set)과 양방향 장단기메모리에 입력되는 데이터 셋 두 가지로 구성된다. <표 2>는 이러한 과정을 통해 구성된 입력 데이터 셋의 예시이다.

3.1.2 단계 2: 예측 단계

단계 2에서는 단계 1에서 생성된 데이터 셋을 바탕으로 모델을 학습시키고 예측을 하게 된다. 단계 2의 밀집층 부분은 주식 가격의 시계열적 측면을 고려하지 않고 20일치 데이터로 구성된 캔들 스틱과 변동성지수 차트로 생성된 이미지 한 장에 집중하여 익일 주가의 등락을 예측하기 때문에 해당 차트의 이미지 특징 잠재 값을 입력으로 받게 된다. 따라서 입력층은 30개의 노드, 은닉층 (hidden layer)은 10개와 2개 노드로 구성된 2개의 층을 지닌다. 단계 2의 어텐션 메커니즘 부분은 하나의 양방향 장단기메모리 층과 하나의 어텐션 층으로 이루어져 있으며, 각 층의 셀은 룩백 (lookback) 기간과 동일한 11개의 셀을 가지고 있다. 이후 어텐션 메커니즘 부분은 2개의 층을 가진 밀집층으로 연결되고 밀집층의 각 노드 개수는 각각 30개와 10개로 구성되어 있다. 이후 밀집층 부분과 어텐션 메커니즘 부분을 결합하여 익일 주가의 등락을

예측한다.

본 연구에서는 익일의 주가를 상승 혹은 하락으로 예측하는 것을 그 목적으로 한다. 이때  $P_t$ 는 당일 증가,  $P_{t+1}$ 는 익일 증가로 익일 증가가 당일의 증가보다 높거나 같으면 1, 낮으면 0으로 설정하였다. 이를 식으로 나타내면 식 (10)과 같다.

$$y = \begin{cases} 1, & P_t \leq P_{t+1} \\ 0, & P_t > P_{t+1} \end{cases} \quad (10)$$

3.2 데이터 수집

본 연구에서는 S&P 500 구성 종목 중 임의로 50개의 기업을 선정해 1993. 07. 01-2019. 07. 31까지 약 25년간의 주식관련 데이터로부터 획득한 시가, 최고가, 최저가, 증가(OHLC: open-high-low-close), 거래량을 바탕으로 작성된 캔들스틱과 변동성지수 차트를 기반으로 데이터를 수집하였다. 주식 관련 데이터는 Investing.com으로부터 개별 기업당 약 6,400일의 데이터, 전체 약 32만 건의 데이터를 수집하였다. 본 연구에서 사용된 데이터는 주가 상승이 전체 날짜의 약 48%, 주가 하락이 약 52%로 구성되어 있다. 수집된 데이터의 예는 다음 <표 3>과 같다.

<표 3> 수집된 데이터

날짜	시가	최고가	최저가	증가	거래량	변동%
2019년 07월 22일	203.65	207.23	203.61	207.22	22.28M	2.29%
2019년 07월 23일	208.46	208.91	207.29	208.84	18.36M	0.78%
2019년 07월 24일	207.67	209.15	207.17	208.67	14.99M	-0.08%
2019년 07월 25일	207.02	209.24	206.73	207.02	13.91M	-0.79%

각 캔들스틱 차트는 Bao *et al.*(2017)의 연구에 따라 20일 데이터로 구성하였으며 장단기메모리의 시간 스텝(time step)은 선행 연구들을(Chen *et al.*, 2016; Persio and Honchar, 2016) 기준으로 30일로 구성하였다. 슬라이딩 윈도우(sliding window)는 1일로 설정하였으며 파이썬 3.6버전과 ‘mpl\_finance’라이브러리를 사용하여 캔들스틱 차트를 구축하였다. 또한 관련 연구(Kim and Kim, 2019)를 기반으로 캔들스틱 차트 하단에 거래량 막대 그래프와 꺾은선 그래프로 구성된 변동성 지수 차트를 추가하여 하나의 이미지로 생성하였다. 이러한 수집 절차에 따라 최종적으로 약 31만 개 가량의 차트를 생성하였다.

### 3.3 평가 방법

연구 모델의 성능 평가를 위해 곡선아래면적(AUC: area under the curve)값을 사용하였다. 곡선아래면적은 분류 문제에 있어 정확도(accuracy)에 비해 클래스 불균형에 보다 견고성(robustness)이 있으며 F1 값에 비해 특정 임계 값(threshold)에 덜 영향을 받기 때문에 본 연구에서는 곡선아래면적을 활용하여 모델 간 비교를 하였다. 곡선아래면적은 수신자 조작 특성(ROC: receiver operating characteristic) 곡선 아래의 면적을 의미하는데 이는 가로축을 특이도(specificity), 세로축을 민감도(sensitive)로 하여 시각화한 그래프를 말한다. 특이도와 민감도는 다음 식 (11)과 식 (12)를 활용하여 구할 수 있다.

$$specificity = \frac{TN}{TN+FP} \quad (11)$$

$$sensitive = \frac{TP}{P} \quad (12)$$

TN: 실제 값이 부정인데 부정으로 예측한 수  
 TP: 실제 값이 긍정인데 긍정으로 예측한 수  
 FP: 실제 값이 부정인데 긍정으로 예측한 수

다만 많은 인공지능 연구에서 모델 간 성능 평

가를 위해 평가 손실과 정확도를 사용하고 있기 때문에 이들 평가 지표를 곡선아래면적과 함께 모델 간 성능 비교를 위한 보조 수단으로 활용하였다.

본 연구에서 제시한 방법론을 간단하게 요약하면 다음과 같다. 우선 주식투자에 사용되는 캔들스틱 차트와 시장 참여자들의 시장 기대치라고 할 수 있는 변동성지수를 수직 결합하여 하나의 차트로 구성한다. 이렇게 구성된 차트를 입력 값으로 변이형 오토인코더를 적용하여 잠재 값을 추출한다. 이때 얻어진 잠재 값은 캔들스틱 차트와 변동성지수 차트의 특성을 나타내는 값이라 할 수 있다. 이러한 잠재 값을 시계열 예측을 위해 어텐션 메커니즘을 적용한 장단기 메모리와 밀집층에 입력 값으로 적용하여 주가가격의 등락을 예측하게 된다. 이러한 예측 값의 성능을 평가하기 위하여 곡선아래면적 값을 도출하고 이를 기반으로 성능을 평가한다.

## IV. 분석 및 결과

### 4.1 실험

본 연구에서 제안한 VATD 모델의 정확한 성능을 파악하기 위해 총 6가지 실험을 진행하였다. 먼저 캔들스틱 차트 기반 예측 모델과 숫자 데이터 기반 예측 모델의 성능 비교를 위해 숫자 데이터에 어텐션 메커니즘이 적용된 양방향 장단기메모리 방식을 적용하여 예측을 시도하였다. 이를 통해 숫자와 이미지 데이터 간의 예측 정확도 비교를 시도하였다. 또한, 수집된 캔들스틱 차트를 기반으로 본 연구에서 제안한 VATD와 합성곱 신경망(CNN), 변이형 오토인코더(VAE), 변이형 오토인코더와 합성곱 신경망 조합(VAE+CNN), 단순 변이형 오토인코더와 어텐션 메커니즘의 단순 조합(VAE+Attention)의 모델 간 성능 비교를 위한 다섯 가지 실험을 진행하였다. <표 4>는 진행된 실험들을 정리한 표이다.

<표 4> 실험 모델

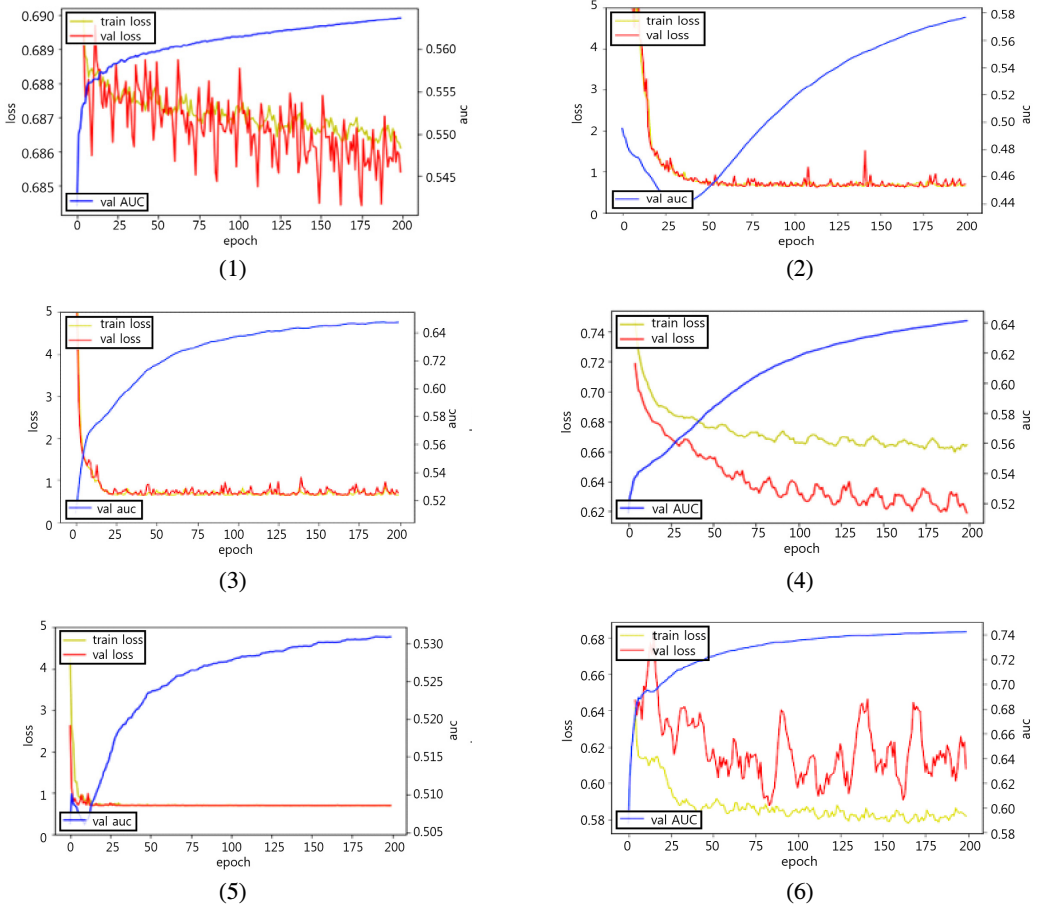
사용 데이터	사용 모형	최적화 알고리즘	비고
주식 데이터(숫자 데이터)  캔들스틱 차트, 변동성지수 차트 (이미지 데이터)	Attention	rmsprop	Attention 방식을 사용하여 예측
	CNN	rmsprop	CNN을 사용하여 예측
	VAE	rmsprop	잠재 값을 히든층으로 예측
	VAE+CNN	rmsprop	VAE와 CNN방식을 결합하여 예측
	VAE+Attention	rmsprop	VAE와 Attention 방식을 결합하여 예측
	VATD(본 연구)	rmsprop	VAE, Attention, Dense layer 결합하여 예측

### 4.2 실험결과 분석

<그림 5>는 앞서 언급한 6개의 딥 러닝 모델들의 실험결과를 그래프로 표현한 것이다. 학습은

파이썬 3.6 버전에서 텐서플로우와 케라스로 진행하였고 모델의 손실이 더 이상 하락하지 않는 200 에포크(epoch) 동안 진행하였다.

<그림 5>의 (1)은 숫자 기반 주식 데이터를 어



<그림 5> 모델 별 평가 손실과 곡선아래면적

텐션 메커니즘으로 학습시킨, (2)는 20일치의 캔들스틱 차트를 합성곱 신경망으로 학습시킨, (3)은 변이형 오토인코더의 잠재 값을 히든층의 입력 값으로 하여 학습시킨 예측결과이다. (4)는 합성곱 신경망과 변이형 오토인코더 방식을 결합하여 학습시킨, (5)는 변이형 오토 인코더의 잠재 값을 어텐션 모델의 입력 값으로 하여 학습시킨, (6)은 본 연구에서 제안한 VATD 모델로 학습시킨 평가 손실과 곡선아래면적 결과이다.

200 에포크 학습 후 각 모델 별 평가 손실과 곡선아래면적(AUC)을 확인한 결과 (1)방식 평가 손실은 0.6863, 정확도는 0.5506, 곡선아래면적은 0.5637로 나타났다. (2)방식의 경우 각각 0.6966, 0.5856, 0.5766, (3)방식의 경우 각각 0.6929, 0.5789, 0.6469, (4)방식의 경우 각각 0.6087, 0.7031, 0.6786, (5)방식의 경우 각각 0.6871, 5359,0.5344, 본 연구에서 제안한 VATD 방법인 (6)방식의 경우 각각 0.5609, 0.7114, 0.7417으로 나타났다. 이를 정리하면 다음 <표 5>와 같다.

<표 5> 모델 별 결과표

	평가손실	정확도	곡선아래면적
Attention	0.6863	0.5506	0.5637
CNN	0.6966	0.5856	0.5766
VAE	0.6929	0.5789	0.6469
VAE, CNN	0.6087	0.7031	0.6786
VAE+Attention	0.6871	0.5359	0.5344
VATD	<b>0.5609</b>	<b>0.7114</b>	<b>0.7417</b>

본 연구에서 사용된 데이터가 주가 상승이 전체 날짜의 약 48%, 주가 하락이 전체 날짜의 약 52%로 구성되어 있는 것으로 볼 때, 검토된 모든 모델들은 좋은 성능을 보인다고 할 수 있다. 특히 본 연구에서 제안한 VATD의 경우 다른 모델에 비해 손실(0.5609)은 가장 낮고, 정확도(0.7114)와 곡선아래면적(0.7417)의 값은 각각 가장 높은 것으로 나타나 다른 모델에 비해 가장 성능이 좋은 것으로 확인되었다.

나아가 본 연구에서 제안한 VATD의 성능을 객

<표 6> 기존연구와 본 연구와의 성능 비교

구분	저자	사용 데이터셋	사용 방법	분석 기간	정확도		
정형	기본	Niaki and Hoseinzade (2013)	S&P 500	Technical indexes	ANN	1994.3~2008.6	51.78%
		Persio and Honchar (2016)	S&P 500	Technical indexes	ANN, WCNN, RNN	2006~2016	56.9%
		Zhong and Enke(2017)	S&P 500	Technical indexes	ANN, PCA	2003.6~2013.5	59.2%
		Song(2018)	S&P 500	Technical indexes	RNN, SVM, XGBoost	2010.1~2017.12	64.33%
비정형	감성	Ding <i>et al.</i> (2015)	S&P 500	Finance news	EB, CNN	2006.10~2013.11	65.48%
		Nguyen <i>et al.</i> (2015)	Twitter, price(OHLC)	Various	Various	2012.7~2013.7	56%
		Huynh <i>et al.</i> (2017)	S&P 500	Reuters News dataset	GRU	2006.10~2013.12	59.98%
	이미지	Guo <i>et al.</i> (2018)	TAIFEX	Candlestick chart	AE, CNN	1998~2016	69.11%
	감성+이미지	본 연구	S&P 500	Candlestick chart, VIX chart	VAE, Attention	1993.7~2019.7	71.14%

관적으로 평가하기 위해 기존 관련 연구의 결과 값들과 비교하였다(<표 6> 참조). Niaki and Hoseinzade (2013), Persio and Honchar(2016), Song(2018), Zhong and Enke(2017)의 연구는 정형적인 주식데이터를 이용하여 익일 주식가격의 등락을 예측한 연구들인 반면 Ding *et al.*(2015), Huynh *et al.*(2017), Nguyen *et al.*(2015)의 연구는 경제관련 뉴스나 트위터의 게시물을 기반으로 제한적으로나마 주식시장의 감성을 고려하여 익일 주식가격의 등락을 예측한 연구들이다. 마지막으로 Guo *et al.*(2018)의 연구는 본 연구와 유사하게 캔들스틱 차트를 활용하여 익일 주식가격의 등락을 예측한 연구이다.

기존 연구들이 성능 평가로 활용한 값이 정확도이기 때문에 기존 연구와의 성능 비교를 위해 정확도를 사용하였다. 비교 결과 정형적인 주식데이터를 활용한 연구 뿐만 아니라 감성, 이미지 등을 활용한 기존 연구에 비해 본 연구가 더 높은 정확도 값을 보인 것을 확인할 수 있었다.

### 4.3 추가 실험

본 연구에서 제안한 VATD 모델의 실용성 확인을 위해 추가 실험을 진행하였다. 2011년 10월 14일부터 2019년 7월 31일까지 기간 동안 뉴욕증권거래소에 상장된 5개 기업의 주식을 무작위로 선정하여 시뮬레이션을 진행하였다. 벤치마크 수익률은 매수 후 보유(buy and hold) 방식을 사용하여 2011년 10월 14일에 주식을 산 후 계속 보유하여 2019년 7월 31일에 판매하는 것으로 가정하였다. 대조군은 본 연구에서 제안한 VATD 모델이 예측하는 등락에 따라 그 날 종가에 주식을 판매하고 구매하는 것으로 설정하였다. 대상 기업은 시뮬레이션 기간 동안 주가가 10% 이상 상승한 기업, 주가가 10% 이상 하락한 기업, 주가의 변동이  $\pm 10\%$  미만인 기업 3가지로 나누어 진행하였다. 대부분의 경우 VATD 모델이 더 우수한 성능을 보였다. 특히 시뮬레이션 기간 중 주가가 하락한 기업들의 경우 본 연구에서 제안한 모델이 더욱

우수한 성능을 보였다. <표 7>은 이러한 시뮬레이션의 결과를 정리한 것이다.

<표 7> 추가 실험 결과

	매수 후 보유	본 연구(VATD)
10% 이상 상승 기업		
H&R Block	86%	<b>170%</b>
Foot Locker	89.8%	<b>230.2%</b>
Dupont	115.9%	<b>167%</b>
10% 이상 하락 기업		
Ford	-17.5%	<b>81%</b>
Jefferies Financial Group	-13.8%	<b>140%</b>
Limited brands	-35.1%	77%
$\pm 10\%$ 미만 기업		
Helmerich & Payne	2.3%	<b>91%</b>
Kellogg	6.1%	<b>122.6%</b>
Newell Brands	8.4%	0.1%

10% 이상 주가가 상승한 기업인 H&R Block, Foot Locker, Dupont의 경우 본 연구의 모형이 비교 대상인 매수 후 보유 방식보다 약 1.5배에서 3배 더 높은 수익을 보였다. 10% 이상 주가가 하락한 기업군에서도 역시 본 연구에서 제안한 모델이 더 좋은 성능을 보였다. 반면 주가의 변동이  $\pm 10\%$  미만인 3개의 기업에서는 혼합된 결과가 나타났다. Helmerich & Payne과 Kellogg의 경우 본 연구에서 제안한 VATD 모델이 더 좋은 수익을 보였으나 Newell brands의 경우 매수 후 보유 방식이 더 높은 수익을 보였다. 이와 같이 변동성이 낮은 기업군에서 혼합된 결과가 나타나는 이유는 개별 기업의 특성으로 인한 것으로 설명할 수 있을 것이다. 예를 들면, 본 연구의 대상인 Newell Brands의 주식가격의 변동은 2018년 파산한 토이저러스(Toys “R” Us)의 여파가 매우 밀접하게 관련되어 있다. 이러한 사실을 고려해 볼 때, 개별 기업의 특정 사건으로 인한 영향은 본 연구에서 제안한 VATD모형이 반응하기 힘든 것으로 판단된다.

## V. 결 론

### 5.1 연구 요약

캔들스틱 차트를 활용하여 주식가격 예측을 시도한 기존 인공지능 연구들은 학습과정에서의 과적합 위험, 캔들스틱 차트의 시계열적 특성을 고려하지 못한다는 점, 시장 참여자들의 감정 상태를 고려하지 못한다는 점 등이 한계로 지적되어 왔다. 본 연구에서는 이러한 문제점을 해결하고자 캔들스틱 차트에 변동성지수 차트를 추가하여 데이터를 생성하고 이를 변이형 오토인코더, 어텐션 메커니즘, 그리고 밀집층을 앙상블하여 분석하는 VATD 모델을 제안하였다. 제안된 모델의 성능 평가를 위해 S&P 500 구성 종목 중 임의로 선정된 50개 기업으로부터 수집된 32만 건의 데이터에 해당 모델을 적용하여 얻어진 평가손실, 정확도, 곡선아래면적 값을 다른 5개의 모델과 비교하여 분석하였다. 분석 결과, 합성곱 신경망, 변이형 오토인코더, 변이형 오토인코더와 합성곱 신경망의 결합 모델 등이 55~64%의 정확도와 0.64이하의 곡선아래면적 값을 보인 반면 본 연구에서 제안하는 VATD 모델은 최대 71%의 정확도와 0.74의 곡선아래면적 값을 보여주어 성능이 가장 우수한 것을 확인하였다 (<표 5> 참조). 또한 기존 주식가격 예측 연구 결과와 비교했을 때에도 더 우수한 성능을 보인 것을 확인하였다 (<표 6> 참조). 추가적으로 2011년 에서 2019년까지의 데이터를 활용한 시뮬레이션 투자 결과 매수 후 보유 전략을 채택한 벤치 마크와 비교하여 본 연구 모델이 더 우수한 성과가 있는 것을 확인하였다 (<표 7> 참조).

### 5.2 연구의 시사점

본 연구는 다음과 같은 점에서 이론적 의의가 있다. 첫째, 시장 참여자들의 감정상태를 고려함으로써 주식가격 예측의 정확성을 개선하였다. 대

중의 감정은 개인의 행동과 의사결정에 중요한 영향을 미치며 이는 주식 시장도 다르지 않음에도 불구하고 기존 캔들스틱 차트 기반 주식가격 예측 연구들은 시장 참여자들의 감정상태를 고려하지 못했다. 본 연구에서는 변동성지수 차트를 추가하여 시장 참여자들의 감정상태를 고려하고자 하였다. 변동성 지수는 시장 참여자들의 시장에 대한 전망을 나타내는 지표로 시장 참여자들이 시장에 대해 어떻게 느끼고 있는지를 일정 정도 반영한다고 할 수 있다(Bekaert and Hoerova, 2014). 따라서 캔들스틱 차트와 변동성지수 차트를 함께 고려함으로써 예측의 정확성을 높일 수 있었다. 특히 본 연구의 추가 실험 결과와 같이 시장 상태에 영향을 크게 받는 업종의 기업이나, 상황이 좋지 않은 기업 등에서 더 좋은 투자 결과를 보인 것으로 판단해 볼 때, 시장 참여자들의 감정을 고려하는 것은 보다 정밀한 주식가격 예측에 크게 도움이 된다고 할 수 있다.

둘째, 본 연구는 기존 연구에 비해 분석 기업의 수와 기간을 늘려 자료의 대표성과 결과의 일반화 가능성을 증진시켰다. 기존 주식가격 예측 연구들은(Ding *et al.*, 2015; Huynh *et al.*, 2017; Nguyen *et al.*, 2015; Niaki and Hoseinzde, 2013) 데이터 셋 구성에 있어 대부분 약 1년에서 10년 가량의 기간을 사용함으로써 경기 동향, 경제 정책 등과 같은 특정 요인에 의한 주식 시장의 움직임을 통제하지 못함으로써 자료의 대표성과 일반화 가능성에 일정 정도 한계가 있었다. 이와 달리 본 연구에서는 데이터 셋을 구성함에 있어 기존 연구에 비해 약 2.5배인 25년 가량의 기간을 활용함으로써 단점을 극복하고자 하였다. 1993년부터 2019년까지를 분석기간으로 잡아 1990년대 후반, 서브 프라임 사태 등 주식 시장에 영향을 미치는 다양한 요인을 고려하게 됨으로써 자료의 대표성과 일반화 가능성을 개선하였다.

셋째, 시계열 기간이 길 경우 발생하는 장기의존성 문제를 해결함으로써 주식가격 변동의 예측 성능을 향상시켰다. 주식가격의 변화처럼 시계열

기간이 길 경우 장기의존성 문제가 발생한다. 그럼에도 불구하고 기존 연구들은 이를 고려하지 않거나 이를 보완하지 않은 모델을 사용함으로써 예측의 정확성이 상대적으로 낮았다. 본 연구에서는 입력 전체를 고려, 특정 부분에 집중하여 결과값을 도출하는 어텐션 메커니즘을 활용함으로써 이러한 단점을 해결하였다. 또한 양방향으로 학습하는 양방향 장단기메모리 알고리즘을 이용하여 25년이라는 장기 시계열 데이터 셋을 효율적으로 활용할 수 있었다. 나아가 예측에 있어 상대적 중요도가 높을 것으로 판단되는 예측 전 20일치의 자료에 좀 더 가중치를 부여하기 위해 밀집층을 추가함으로써 이러한 문제를 해결하고자 하였다.

본 연구는 또한 다음과 같은 점에서 실무적 의의가 있다. 첫째, 주식가격 예측을 위한 특징 추출에 변이형 오토인코더를 사용함으로써 차원 축소 효과를 얻을 수 있었다. 본 연구의 데이터 셋은 시가, 종가, 최고가, 최저가, 거래량 뿐만 아니라 변동성 지수가 추가되어 구축되었다. 변동성지수의 추가는 시장 참여자들의 감정 상태를 고려할 수 있다는 장점이 있는 반면 고려할 변수가 많아짐에 따라 새로운 데이터에 적용하기 어려워지는 것과 같은 문제점이 발생할 수 있다. 예를 들면, 본 연구와 같이 시가, 종가, 최고가, 최저가, 거래량, 변동성지수를 포함한 데이터 셋 분석을 위해 캔들스틱 차트와 변동성지수 차트를 기반으로 20일치 그래프를 구성하기 위해서는 140개의 차원이 필요하다. 이를 장단기메모리 모형으로 분석하기 위해서 30일의 기간을 고려한다면 140개의 차원으로 구성된 11개의 데이터 셋이 수평으로 연결되어야 한다. 이와 같이 고려할 변수가 많아지면 많아질 수록 데이터 셋의 용량이나 모델의 성능에 큰 영향을 미쳐 성능이 저하되는 문제가 발생할 수 있다. 본 연구에서는 20일치 이미지를 30개 차원으로 축소하여 데이터 셋을 구축함에 따라 컴퓨팅 파워의 절약 뿐만 아니라 고차원 데이터 셋에서 발생할 수 있는 잠재적인 문제점을 해결하였다. 따라서 본 연구에서 제안한 방법을 활용함으로써 실무자

들은 보다 적은 컴퓨팅 파워를 활용하여 고차원 데이터 셋을 분석할 수 있을 것이다. 둘째, 실무자들이 활용할 수 있는 투자 상담을 위한 하나의 대안을 제시하였다. 우리나라가 고령화 사회에 진입하면서 안정성이 보장된 투자에 대한 관심이 높아지고 있다(김기호, 유경원, 2008). 이에 본 연구에서는 기존 주식가격 등락 예측에 사용되던 캔들스틱 차트와 함께 변동성 지수를 추가로 고려함으로써 기존의 연구들 보다 더 나은 성능의 모델을 개발하였다. 주식가격의 정확한 예측을 원하는 실무자들의 경우 본 연구에서 제안한 VATD 모델을 기존 투자 모형과 함께 고려함으로써 안정적인 주식 투자 및 투자자 상담을 위한 보다 다양한 대안을 확보할 수 있을 것으로 기대한다.

### 5.3 한계점 및 향후 연구 방향

본 연구는 다음과 같은 한계점이 있으며 향후 연구를 통해 이를 해결해야 할 것이다. 첫째, 본 연구에서는 데이터 학습 시 주식의 매수와 매도에서 발생하는 거래비용을 고려하지 않았기 때문에 실무에 적용함에 있어 주의가 필요하다. 향후 연구에서는 모델 학습에 있어 주식 거래 시 발생하는 거래 비용을 고려함으로써 보다 실용적이고 정확한 주식가격 예측 모델을 구축할 수 있을 것으로 기대한다. 둘째, 본 연구에서는 캔들스틱 차트에서 제공한 지표 이외에 기업의 특성을 나타낼 수 있는 주가수익률(PER: price earning ratio)이나 주가순자산비율(PBR: price to book-value ratio) 등과 같은 다른 지표를 고려하지 못했다. 본 연구의 추가 실험 부문에서 언급한 바와 같이 제안된 VATD 모델은 각 기업별 특이사항을 반영하지 못하기 때문에 파업, 법적 소송 등 특정 기업만이 가지는 사건이나 상황의 변동 등에서 발생하는 위험인 비체계적 위험(unsystematic risk)에 노출된 기업의 주식가격 예측이 정확하지 않을 가능성이 있다. 따라서 시장 전체 상황보다 기업에 대한 개별 의견에 집중하는 감성 분석 기반 주가예측 방



법을 추가함으로써 VATD 모델의 단점을 개선할 수 있을 것으로 기대한다. 마지막으로 본 연구는 모든 차트를 하나의 이미지로 합쳐 분석을 실시하였다. 그러나 최근 여러 장의 차트를 개별적으로 입력하여 분석함으로써 예측의 성능을 개선할 수 있다는 주장이 제기되고 있다(Ham et al., 2019). 따라서 각 차트를 개별적으로 입력 값으로 하여 분석함으로써 보다 좋은 결과를 기대할 수 있는 가능성이 있다.

## 참 고 문 헌

- [1] 김기호, 유경원, “인구고령화가 인적자본 투자 및 금융시장에 미치는 영향”, *보험개발연구*, 제19권, 제3호, 2008, pp. 165-207.
- [2] 김현모, 박재홍, “온라인주식게시판 정보가주식투자자의 거래행태에 미치는 영향”, *Information Systems Review*, 제18권, 제2호, 2016, pp. 23-38.
- [3] 신동하, 최광호, 김창복, “RNN과 LSTM을 이용한 주가 예측을 향상을 위한 딥러닝 모델”, *한국정보기술학회논문지*, 제15권, 제10호, 2017, pp. 9-16
- [4] Bahdanau, D., K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate”, *3rd International Conference on Learning Representations, ICLR*, 2015, Available at <http://arxiv.org/abs/1409.0473>.
- [5] Bao, W., J. Yue, and Y. Rao, “A deep learning framework for financial time series using stacked autoencoders and long-short term memory”, *PLOS ONE*, Vol.12, No.7, 2017, e0180944.
- [6] Bekaert, G. and M. Hoerova, “The VIX, the variance premium and stock market volatility”, *Journal of Econometrics*, Vol.183, No.2, 2014, pp. 181-192.
- [7] Bengio, Y., P. Frasconi, and P. Simard, “The problem of learning long-term dependencies in recurrent networks”, *IEEE International Conference on Neural Networks*, 1993, pp. 1183-1188.
- [8] Braun, H. and J. S. Chandler, “Predicting stock market behavior through rule induction: Anthe-learning-from-example-approach”, *Decision Science*, Vol.18, No.3, 1987, pp. 415-429.
- [9] Chen, J. F., W. L. Chen, C. P. Huang, S. H. Huang, and A. P. Chen, “Financial time-series data analysis using deep convolutional neural networks”, *7th International Conference on Cloud Computing and Big Data (CCBD)*, 2016, pp. 87-92.
- [10] Chu, W. and D. Cai, “Stacked similarity-aware autoencoders”, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*, 2017, pp. 1561-1567
- [11] Chung, H. and K. S. Shin, “Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction”, *Neural Computing and Applications*, Vol.32, 2020, pp. 7897-7914.
- [12] Ding, G. and L. Qin, “Study on the prediction of stock price based on the associated network model of LSTM”, *International Journal of Machine Learning and Cybernetics*, Vol.11, 2020, pp. 1307-1317.
- [13] Ding, X. Y. Zhang, T. Liu, and J. Duan, “Deep learning for event-driven stock prediction”, *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*, 2015, pp. 2327-2333.
- [14] Fama, E. F., “The behavior of stock-market prices”, *The Journal of Business*, Vol.38, No.1, 1965, pp. 34-105.
- [15] Graves, A. and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures”, *Neural Networks*, Vol.18, No.5, 2005, pp. 602-610.

- [16] Guo, S. J., F. C. Hsu, and C. C. Hung, “Deep candlestick predictor: A framework toward forecasting the price movement from candlestick charts”, *9th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP)*, 2018, pp. 219-226.
- [17] Ham, Y., J. Kim, and J. Luo, “Deep learning for multi-year ENSO forecasts”, *Nature*, Vol.573, 2019, pp. 568-572.
- [18] Hochreiter, S. and J. Schmidhuber, “Long short-term memory”, *Neural Computation*, Vol.9, No.8, 1997, pp. 1735-1780.
- [19] Huynh, H. D., L. M. Dang, and D. Duong, “A new model for stock price movements prediction using deep neural network”, *Proceedings of the Eighth International Symposium on Information and Communication Technology - SoICT*, 2017, pp.57-62.
- [20] Karevan, Z. and J. Suykens, “Transductive LSTM for time-series prediction: An application to weather forecasting”, *Neural Networks*, Vol.125, 2020, pp. 1-9.
- [21] Kim, T. and H. Y. Kim, “Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data”, *PLOS ONE*, Vol.14, No.2, 2019, e0212320.
- [22] Kingma, D. P. and M. Welling, “Stochastic gradient VB and the variational auto-encoder”, *Second International Conference on Learning Representations*, 2013, pp. 1-14.
- [23] Liu Y., Z. Qin, P. Li, and T. Wan, “Stock volatility prediction using recurrent neural networks with sentiment analysis”, *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, 2017, pp. 192-201
- [24] Liu, H. and B. Song, “Stock price trend prediction model based on deep residual network and stock price graph”, *11th International Symposium on Computational Intelligence and Design (ISCID)*, 2018, pp. 328-331.
- [25] Marshall, B. R., Young, M. R., and L. C. Rose, “Candlestick technical trading strategies: Can they create value for investors?”, *Journal of Banking & Finance*, Vol.30, No.8, 2006, pp. 2303-2323.
- [26] Murphy, J. J., *Intermarket Analysis: Profiting from Global Market Relationships*, John Wiley & Sons, Toronto, 2011.
- [27] Nguyen, T. H., K. Shirai, and J. Velcin, “Sentiment analysis on social media for stock movement prediction”, *Expert Systems with Applications*, Vol.42, No.24, 2015, pp. 9603-9611.
- [28] Niaki, S. T. A. and S. Hoseinzade, “Forecasting S&P 500 index using artificial neural networks and design of experiments”, *Journal of Industrial Engineering International*, Vol.9, No.1, 2013, pp. 1-9.
- [29] Ou, J. A. and S. H. Penman, “Accounting measurement, price earnings ratio, and the information-content of security prices”, *Journal of Accounting Research*, Vol.27, 1989, pp. 111-144.
- [30] Pagolu, V. S., K. N. R. Challa, G. Panda, and B. Majhi, “Sentiment analysis of twitter data for predicting stock market movements”, *International conference on Signal Processing, Communication, Power and Embedded System (SCOPEs)*, 2016, pp. 1-5.
- [31] Patalay, S. and M. R. Bandlamudib, “Stock price prediction and portfolio selection using artificial intelligence”, *Asia Pacific Journal of Information Systems*, Vol.30, No.1, 2020, pp. 31-52
- [32] Persio, L. D. and O. Honchar, “Artificial neural networks architectures for stock price prediction: Comparisons and applications”, *International Journal of Circuits, Systems and Signal Processing*, Vol.10, 2016, pp. 403-413.

- [33] Rush, A. M., S. Chopra, and J. Weston, "A neural attention model for abstractive sentence summarization", *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2015, pp. 379-389.
- [34] Schuster, M. and K. K. Paliwal, "Bidirectional recurrent neural networks", *IEEE Transactions on Signal Processing*, Vol.45, No.11, 1997, pp. 2673-2681.
- [35] Song, Y., Stock Trend Prediction: Based on Machine Learning Methods, Master Thesis, UCLA, 2018, Available at <https://escholarship.org/uc/item/0cp1x8th>.
- [36] Yao, K., G. Zweig, and B. Peng, "Attention with intention for a neural network conversation model", NIPS Workshop on Machine Learning for Spoken Language Understanding and Interaction 2015, 2015, Available at <http://arxiv.org/abs/1510.08565>.
- [37] Zhipeng, J. and L. Chao, "Financial time series forecasting based on characterized candlestick and the support vector classification with cooperative coevolution", *Journal of Computers*, Vol.14, No.3, 2019, pp. 195-209.
- [38] Zhong, X. and D. Enke, "Forecasting daily stock market return using dimensionality reduction", *Expert Systems with Applications*, Vol.67, 2017, pp. 126-139.

## Chart-based Stock Price Prediction by Combing Variation Autoencoder and Attention Mechanisms

Sanghyun Bae\* · Byounggu Choi\*\*

### Abstract

Recently, many studies have been conducted to increase the accuracy of stock price prediction by analyzing candlestick charts using artificial intelligence techniques. However, these studies failed to consider the time-series characteristics of candlestick charts and to take into account the emotional state of market participants in data learning for stock price prediction. In order to overcome these limitations, this study produced input data by combining volatility index and candlestick charts to consider the emotional state of market participants, and used the data as input for a new method proposed on the basis of combining variation autoencoder (VAE) and attention mechanisms for considering the time-series characteristics of candlestick chart. Fifty firms were randomly selected from the S&P 500 index and their stock prices were predicted to evaluate the performance of the method compared with existing ones such as convolutional neural network (CNN) or long-short term memory (LSTM). The results indicated the method proposed in this study showed superior performance compared to the existing ones. This study implied that the accuracy of stock price prediction could be improved by considering the emotional state of market participants and the time-series characteristics of the candlestick chart.

**Keywords:** *Candlestick chart, Variation autoencoder, Attention mechanism, Volatility index, Stock price prediction*

---

\* Assistant Manager, KCNET

\*\* Corresponding Author, Professor, College of Business Administration, Kookmin University

## ◎ 저 자 소 개 ◎



**배 상 현 (baesang0522@gmail.com)**

현재 KCNET 데이터 사업부에 대리로 재직 중이다. 국민대학교 데이터사이언스 석사를 취득하였다. 주요 연구분야는 데이터사이언스, 딥러닝 등이며 경영정보학회 학술대회에서 발표를 하였다.



**최 병 구 (h2choi@kookmin.ac.kr)**

현재 국민대학교 경영대학 경영학부 교수로 재직 중이다. KAIST경영공학 석사 및 박사학위를 취득하였다. 국민대학교에 부임하기 전에는 University of Sydney, School of Information Technologies에서 조교수로 재직하였다. 주요 연구분야는 지식경영, 소셜미디어 어널리틱스, 데이터사이언스 등이며 지금까지 이와 관련하여 Journal of Association for the Information Systems, Journal of MIS, IEEE Transactions on Engineering Management, I&M, APJIS, 지식경영연구 등을 포함한 다수의 국내 외 학술지에 논문을 게재하였다.

논문접수일 : 2020년 12월 07일

게재확정일 : 2020년 12월 23일

1차 수정일 : 2020년 12월 21일