빅데이터 기반 가상화폐 공시 분석 AI 연구

정용빈, 안상훈, 이세욱, 박현성, 이도윤, 이효경, 안정운, 고진광
순천대학교 컴퓨터공학과
e mail : fkgeek@naver.com, yu10978772@gmail.com, sjskajrj123@naver.com,
shpin1209@gmail.com, ldodoril@naver.com, hyo961218@gmail.com, ajy1100@gmail.com,
kjg@scnu.ac.kr

A Study On AI Machine BigData based Cryptocurrency News

Yong-Bin Jeong, Sang-Hun An, Se-Uk Lee, Hyun-Sung Park,
Do-Yun Lee, Hyo-Kyung Lee, Jeong-Wun An, Jin-Gwang Koh
Dept. of Computer Engineering, Sunchon National University

요 약

인터넷 가상화폐 거래소에 게시되는 공시를 분석하여 향후 게시된 공시가 가상화폐 가격에 줄 영향을 예측해보고자 한다. 과거 게시된 공시 데이터를 바탕으로 딥러닝을 이용해 영향치를 도출해낸다. 프로그램은 Python 언어로 작성하였으며, django 웹 프레임워크를 이용하여 결과치를 보여줄 수 있는 웹 사이트를 구현하였다. 또한 공시를 통합한 데이터를 이용하여 새로 게시되는 공시가 해당 가상화폐의 가격에 줄 예측치를 분석하였으며 해당 공시가 가상화폐의 가격에 어떻게 영향을 줄지 쉽게 판단할 수 없는 사용자들에게 도움을 줄 수 있을 것이다.

1. 서론

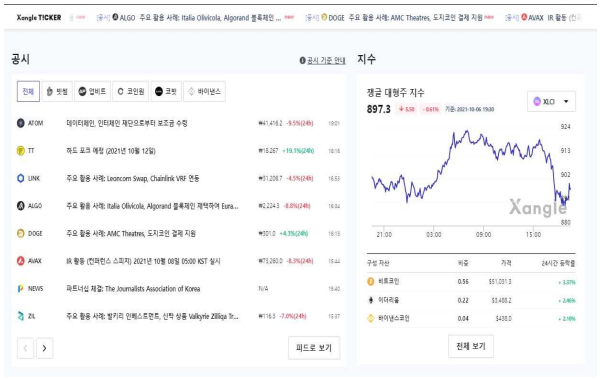
기존의 화폐와는 달라 생소했던 가상화폐는 이제 사람들의 생활 속에 깊이 들어오게 되었다. 지금도 가상화폐에 대한 관심이 꾸준히 증가하고 있으며, 여러 가상화폐 거래소에서 거래되고 있다. 가상화폐의 가치는 가상화폐에 관한 주요 이슈 즉, '공시'에 의해 결정되고, 여러 거래소에 게시되어 빠르게 영향을 주게 된다. 2020년 5월, 비트코인의 공급량이 조절되면서 가상화폐의 접근성이 개선되며 막대한 금액을 운용하는 헤지펀드나 자산운용사, 기관 투자자까지 가상화폐 시장에 뛰어들고 있다. 기업자금 유입과 그에 더해 잦은 미디어의 가상화폐 관련 콘텐츠 노출로 인한 일반 대중들의 유입 자금까지 더해져 궁극적으로 코인 시장에 천문학적인 액수가 유입되어 오늘날 가상화폐 시장의 규모는 거대해진 상황이다. 국내에서는 특정금융정보법의 시행으로 가상화폐 시장이 안전한가에 대한 걱정도 있지만, 미국과 일본 등 선진국에서 가상화폐 시장의 지속적인 지원과 디지털 화폐(CBDC)의 개발 중요성이 나날이 커지며 긍정적인 소식이 쏟아져 나오고 있다. 최근에는 미국연방준비제도(FED) 파월 의장 또한 디지털 화폐를 직접 언급하는 등 제도권에 편입됨에 따라 더욱 가상화폐 시장의 발전 가능성이 돋보이는 추세다.

본 연구에서는 딥러닝 기술을 이용해 가상화폐에 관한 주요 공시를 분석하여 추후 가상화폐의 가치에 미칠 영향치를 도출하는 방법에 대해 연구하였다.

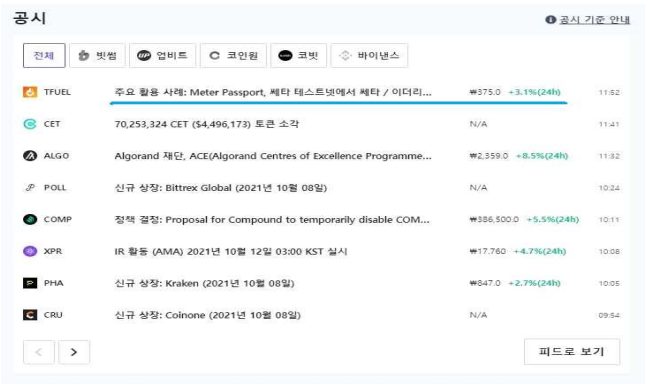
2. 본론

2.1 데이터 수집 및 전처리

데이터를 학습하고 결과치를 도출하기 위한 프로그램 개발에는 Python 프로그래밍 언어를 사용하였으며 머신러닝, 딥러닝 프로그램 작성에 tensorflow 소프트웨어를 사용하였다. 학습시키기 위한 데이터를 수집하기 위해 웹 크롤링을 이용하여 공시 전문 사이트인 쟁글(xangle)에 게시된 공시에서 제목 및 코인 심볼 데이터를 수집하였으며 수집된 데이터에서 해당 공시의 제목이 포함하는 단어 중에서 현재 가상화폐의 가격의 상승에 영향을 줄 수 있는 '신규상장', '파트너쉽', '에어드랍' 등의 단어를 포함하였을 경우 정수 1, '상장취소', '상장폐지' 등 가상화폐의 가격의 하락에 영향을 줄 수 있는 단어를 포함하였을 경우 정수 0을 할당하고, 수집한 데이터들과 하나의 .csv 파일로 통합하여 프로그램 학습에 이용하였다.[그림 1]



[그림 1] 공시 정보 포털 사이트 'xangle'



[그림 3] 같은 날짜의 공시 사이트

2.2 딥러닝 분석 모델 선정 및 구축, 실험

데이터를 분석할 딥러닝 분석 모델로 LSTM 모델을 선정하여 하이퍼 파라미터 튜닝 과정을 거치고 해당 LSTM 모델에 호재와 악재를 구분할 수 있는 분석 알고리즘을 구현하여 적용하여 학습하였다. 프로그램은 수집한 데이터를 통해 학습하고, 공시 사이트에 게시된 공시의 제목을 받아 분석한 후, 그 결과를 출력하게 된다. 공시의 제목을 분석하여 가치 상승에 영향을 줄 경우 100%, 가치 하락에 영향을 줄 경우 0% 에 가깝게 결과값을 도출한다.

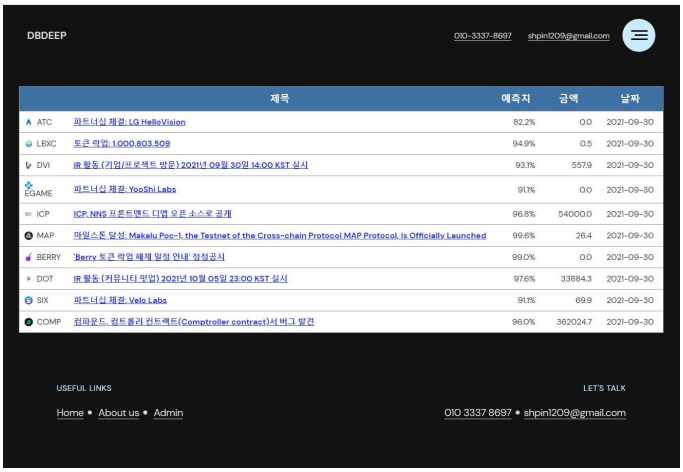
딥러닝 모델을 사용하여 예측한 결과 실제 공시 사이트와도 어느 정도 유사한 것을 확인할 수 있었다. [그림 2] 와 [그림 3] 에 표시된 같은 공시에서 딥러닝 모델은 99.9%로 예측하였으며, 실제로 공시 사이트에서도 해당 가상화폐의 가격이 상승한 것으로 표시되고 있으며, 예측치가 높은 다른 공시들도 대부분 가격의 상승을 보이고 있다.

	제목	예측치	금액	날짜
TFUEL	주요 활용 사례: Meter Passport, 베타 테스트넷에서 베타 / 이더리움 브리지 출시	99.9%	374.6	2021-10-08
CET	70,253,324 CET ($4,496,173) 토큰 소각	92.2%	0.0	2021-10-08
ALGO	Algorand 재단, ACE(Algorand Centres of Excellence Programme) 출시	98.6%	2323.5	2021-10-08
POLL	신규 상장: Bittrex Global (2021년 10월 08일)	100.0%	0.0	2021-10-08
COMP	정책 결정: 컴파운드, COMP 보상 지급을 일시적으로 비활성화하는 것에 대한 제안 통과되었지만 폴연 취소	99.8%	386500.0	2021-10-08
XPR	IR 활동 (AMA) 2021년 10월 12일 03:00 KST 실시	99.5%	17.8	2021-10-08
PHA	신규 상장: Kraken (2021년 10월 08일)	100.0%	847.0	2021-10-08
CRU	신규 상장: Coinone (2021년 10월 08일)	100.0%	0.0	2021-10-08
ADMON	토큰 분배 (20, 2021)	75.6%	0.0	2021-10-08
HBAR	신규 상장: Bitstamp (2021년 10월 12일)	100.0%	452.7	2021-10-08

[그림 2] 딥러닝 모델을 이용한 예측 결과

2.3 예측치를 보여줄 웹 사이트 구현

도출한 데이터를 보여주기 위해 웹 사이트를 제작하였으며, Python 언어에서 제공되는 웹 프레임워크인 장고(django)를 이용하여 구현하였다. 웹 페이지에서는 공시가 계속 업데이트 되면서 공시의 제목과 날짜, 그리고 공시가 가상화폐의 가격에 미칠 영향력을 예측한 예측치를 보여주며, 공시 제목을 클릭하면 해당 공시를 바로 볼 수 있도록 하였다.



[그림 4] 장고(django)로 구현된 웹 페이지

[그림 5] 제목을 클릭했을 때 볼 수 있는 공시

3. 결론

본 연구에서는 과거 거래소에 게시된 공시를 통합한 데이터를 이용하여 새로 게시되는 공시가 해당 가상화폐의 가격에 줄 예측치를 분석하였으며, 분석한 예측치를 보여줄 수 있는 웹페이지를 구현하였다. 본 연구에서의 결과물은 공시를 접하고도 해당 공시가 가상화폐의 가격에 어떻게 영향을 줄지 쉽게 판단할 수 없는 사용자들에게 도움을 줄 수 있을 것이라고 생각한다. 새로운 공시는 하루에도 여러 건 게시되고 있으며, 이 공시에 따라 가상화폐의 가격은 천차만별로 변하게 되기도 한다. 만약 공시가 게시되는 동시에 분석한 예측치를 바로 사용자에게 보여줄 수 있다면, 사용자의 판단에 도움을 줄 수 있을 것이다.

향후 연구에서는 우선 분석의 목표치를 더욱 높이는 것이 목표이다. 학습 알고리즘의 개선을 통해 사용자에게 더 정확한 예측치를 제공할 수 있도록 하는 것이다. 또한, 사용자가 공시에 대한 예측치를 보고 직접 거래를 하는 것이 아니라 사용자가 거래를 원하는 가상화폐에 대한 공시가 게시되었을 때 프로그램이 사용자에게 해당 가상화폐의 거래를 진행할 것인지에 대한 여부를 알림을 통해 물어보고, 거래를 바로 진행하는 방식을 구현할 수 있도록 연구 및 개발을 진행할 예정이다.

참고문헌

[1] https://xangle.io/

[2] Sang-Soog Lee Inhyeok Yoo jinhee Kim “An analysis of public perception on Artificial Intelligence(AI) education using Big Data: Based on News articles and Twitter “

[3] Junsik Lee ,Keon-Woo Kim , Do-Hyung Park ”Empirical Analysis on Bitcoin Price Change by Consumer“, Industry and Macro-Economy Variabl(2018)

[4] Hong June Gi ”Analysis of Sales Volume by Products According to Temperature Change Using Big Data Analysis“(2019)

[5] Mu Moung Cho Han, Yangsok Kim, Choong Kwon Le (Analysis of the Relations between Social Issues and Prices Using Text Mining - Avian Influenza and Egg Prices -)

[6] TaeSun Im , ByungJin Yim “An Empirical Study on the Mutual Influence of the Bitcoin Price and the Financial Market”(2021)

[7] Sung Wook Yoon “Analyzing Impact of Bitcoin Features to Bitcoin Price via Machine Learning Techniques”(2020)