연구결과보고서

2021313030 소프트웨어학과 이재준

**Ⅰ. 연구주제**

미국 기업인 일론 머스크의 개인 트위터의 트윗과 암호화폐 중 하나인 도지코인의 가격 간의 관계성

**Ⅱ. 연구배경**

일론 머스크를 검색하면 연관검색어에 트위터, 도지코인, 비트코인이 나올 정도로 그와 트위터, 암호화폐는 관계가 깊다. 테슬라의 최고경영자인 그가 테슬라 차량구매 때 암호화폐를 거래수단으로 사용한다고 선언하였을 때 암호화폐들의 가격이 큰 폭으로 증가하였다. 또한 그가 암호화폐 결제를 중단한다고 선언하였을 때에는 암호화폐들의 가격들이 큰 폭으로 하락하였다. 그동안 그의 한 마디에 암호화폐의 가격이 오르고 내리는 것을 통해 그가 암호화폐 시장에서 가장 큰 영향력을 가진 인물이라는 것은 누구나 알 수 있는 사실이다.

텍스트이(가) 표시된 사진

자동 생성된 설명

심지어 이러한 그거 암호화폐에 관련되어 긍정적인 내용을 트위터에 올리면 암호화폐 가격이 급등하고, 부정적인 내용을 게시하면 가격이 폭락하는 등 그의 개인 SNS인 트위터의 영향력이 크다는 기사가 쏟아져 나오는 상황이다. 일론 머스크가 가상화폐 시장을 교란한다는 의견도 커지는 가운데 공식 발표가 아닌 트윗의 영향력이 커봤자 얼마나 클 것인가라는 의문과 더불어 가상화폐의 가치의 변동성은 없는 상황에서 개인의 sns에 휘둘리는 일부 투자자들이 잘못한 것이 아닌 가하는 의문이 들었다. 따라서 일론 머스크의 트윗의 내용에 따라 가상화폐의 변동성이 어떻게 변화하는지 직접 분석해보기로 하였다. 가상화폐는 종류가 다양하기 때문에 일론 머스크와 가장 연관이 있는 가상화폐로 알려진 도지코인을 기준으로 연구를 진행하기로 하였다.

**Ⅲ. 연구 수행 방법**

연구는 두 가지 방법으로 진행하였다. 첫 번째 방법은 도지코인의 가격등락률의 절댓값을 구한 뒤, 값이 큰 상위 10개의 자료에 해당하는 날짜를 구해 그날 일론 머스크의 트윗이 있었는 지와 있다면 어떤 내용인지를 살펴보는 것이다. 일론 머스크의 트윗 내용에 따라 가격이 상승하기도 하락하기도 하였다고 하기에 가격이 상승했는지, 하락했는지 여부는 구하지 않고 그 변동 폭을 구하기로 하였다. 변동 폭이 큰 날들에 일론 머스크의 트윗 내용이 가상화폐와 관련이 있다면 도지코인에 대해 일론 머스크의 트윗의 영향력이 상당함을 판단할 수 있을 것이다.

두 번째 방법은 도지코인의 가격과 일론 머스크가 트위터에서 Doze와 Coin이라는 키워드를 언급한 횟수를 나타낸 일별 그래프를 그려 두 그래프의 유사도를 판단할 것이다. 도지코인의 가격이 가파르게 변동하는 추세를 보이는 기간에 일론 머스크가 트윗에서 해당 키워드들을 많이 언급했었다면 그의 트윗의 영향력이 크다고 판단할 수 있을 것이다.

도지코인에 대한 데이터는 2017년 6월부터 2021년 7월까지의 일별로 도지코인의 마감 가격, 고점 가격, 저점 가격, 가격등락률 등을 나타낸 kaggle의 데이터를 사용하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

(<https://www.kaggle.com/tarandeep97/dogecoin-historical-data20172021>)

테이블이(가) 표시된 사진

자동 생성된 설명

(도지코인 데이터셋)

도지코인 데이터를 기준으로 2017년 6월부터 2021년 7월의 일론 머스크의 트윗에 대한 데이터를 찾으려 하였으나 마땅한 데이터셋을 찾지 못하여 직접 일론 머스크의 트윗을 크롤링하여 데이터셋을 만들기로 하였다. 일반적인 크롤링으로 방법으로는 트위터에서 서버 과부하를 방지하기 위해 1주일 이상 지난 트윗을 크롤링할 수 없게 하였기 때문에 파이썬의 snscrape 모듈을 사용하여 크롤링을 진행하였다. 2017년 6월 3일부터 2021년 7월 1일까지의 트윗의 날짜와 내용을 크롤링하여 CSV파일로 저장하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

인코딩을 utf-8규격으로 진행한 결과 트윗 내용의 일부가 깨지는 현상이 발생하였다. 따라서 메모장을 중계 수단으로 활용하여 ANSI 형식으로 인코딩 형식을 바꾸어 문제를 해결하였다.

테이블이(가) 표시된 사진

자동 생성된 설명

(일론머스크 트윗 데이터셋)

**Ⅳ. 연구 결과**

ⅰ. 도지코인의 상위 등락률 날짜와 일론 머스크 트윗과의 관계 분석

도지코인의 등락률의 절댓값의 상위 10개의 값을 추출하기 위해서 도지코인 데이터셋의 Change % 열을 이용하였다. 등락률 값이 아닌 순위가 중요하기 때문에 비교가 쉽도록 % 기호, 천 단위기호 , 를 제거하였고 절댓값을 취하기 위해 -기호도 제거한 뒤 숫자형 데이터로 변환시켰다. 그 후 내림차순 정렬하여 상위 10개의 값을 뽑고 그 행에 해당하는 Date 값을 가져와 날짜형식으로 변환하였다. 그 결과는 다음과 같다.

텍스트이(가) 표시된 사진

자동 생성된 설명

(topTenDateList의 앞 부분에 위치한 날짜부터가 가장 큰 등락률을 보인 날짜이다.)

그 다음으로 tweetData의 date열을 2020-01-15과 같은 도지코인 데이터셋의 날짜데이터와 비교하기 위하여 날짜형 데이터로 바꾼 후 날짜 형식을 변경하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

topTenDateList에 해당하는 날짜에 일론 머스크가 트윗을 올렸다면 그 트윗을 모두 가져오기로 하였다.

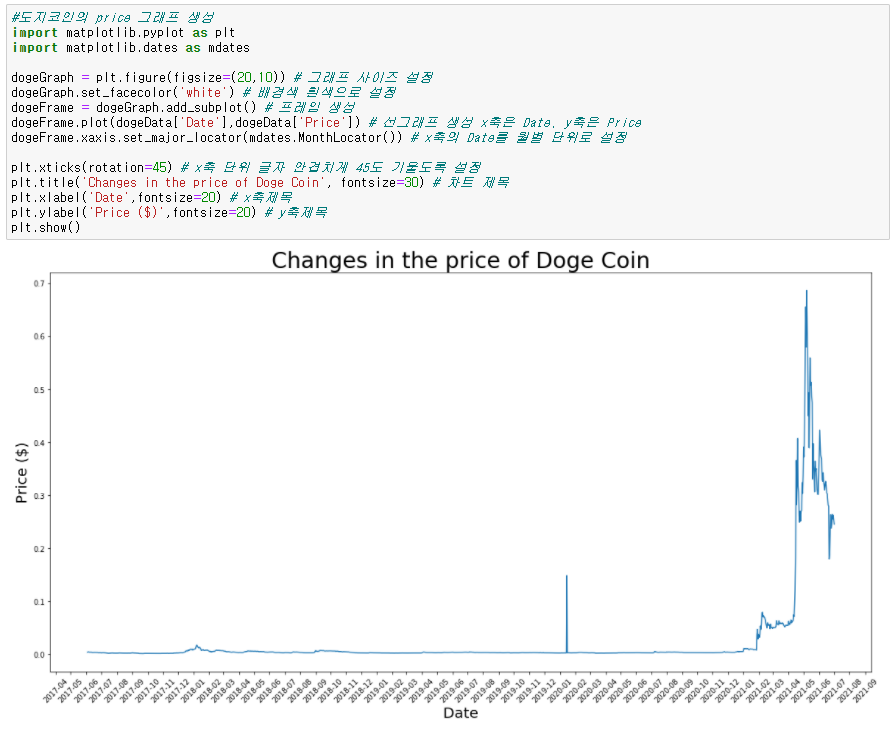
텍스트이(가) 표시된 사진

자동 생성된 설명

그 결과 상위 10개의 등락률 중 10위에 해당하는 날에만 Doge라고 직접적으로 언급한 것을 확인할 수 있었다.

ⅱ. 도지코인 가격 그래프와 일론 머스크의 키워드 언급 트윗 수 그래프 비교

먼저 도지코인의 그래프를 생성하였다. 그래프의 x축 단위는 월별 날짜로 y축은 도지코인의 가격($)를 기준으로 하였다. 그래프는 다음과 같다.

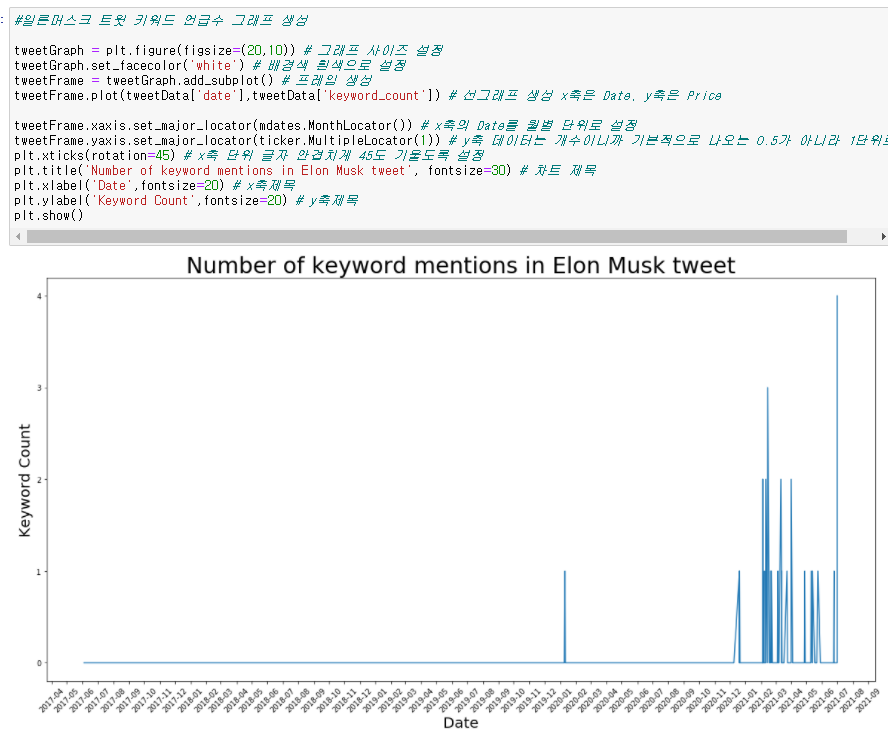


그 다음에는 일론머스크 트윗에서 키워드가 몇 번 나왔는 지를 세어 새로운 열을 생성하였다. 키워드는 Doze, doze, coin, Coin으로 설정하였다. bitcoin은 coin이라는 단어가 안에 포함되어 있어 중복되어 세어질 수 있기 때문에 제외하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

그 뒤 그래프의 x축을 일별 날짜로 y는 키워드 언급 횟수로 설정하여 일론 머스크의 트윗 언급횟수 그래프를 그린 결과는 다음과 같다.



2020년 1월에 일론 머스크가 키워드를 포함한 트윗을 올린 적이 있다고 나타나는데, 2020년 1월에 도지코인의 가격이 급상승하는 그래프를 확인할 수 있었다. 그리고 또한 2021년 1월부터 도지코인의 가격이 큰 폭으로 상승 그래프를 그렸다. 그래프에 따르면 2021년도 1월부터 그래프가 많이 겹쳐지게 나오는데 x축이 일별이므로 많이 겹쳐져 보일수록 그 달에 많이 언급했다고 판별할 수 있다. 따라서 일론 머스크의 트윗의 키워드가 2021년도 1월말부터 급격하게 많이 나타난다는 것을 알 수 있었다.

**Ⅴ. 연구주제에 대한 연구결과 종합 및 결론**

첫 번째 연구에서는 등락률 절댓값 상위 10개에 해당하는 날짜 중 10위에 해당하는 날짜에 Doge와 같이 직접적으로 언급한 것을 확인할 수 있었다. 이는 등락률과 일론 머스크의 트윗이 크게 관련이 없어 보인다고 판별할 수도 있지만, 일론 머스크의 트윗에는 간접적으로 가상화폐 관련해서 뉘앙스를 풍기거나, 내용과는 관계없이 해시태그로 가상화폐를 태그한 경우가 있다는 것을 확인할 수 있었다. 또한 이모티콘을 활용하여 제대로 된 데이터를 얻지못한 트윗 또한 존재하였다.

텍스트이(가) 표시된 사진

자동 생성된 설명

등락률 절댓값 3위에 해당하는 트윗이었던 ‘Make life multiplanetary!’라는 트윗을 직접 확인해보면 내용 자체에는 직접적으로 언급한 것이 없지만 그의 트윗 댓글에서는 doge에 대한 내용이 가득했었다. 이를 통해 직접적으로 트윗으로 언급하지 않았더라도 그날의 그의 행동이라 그동안의 트윗들의 뉘앙스로 간접적으로 영향을 끼친 경우가 많다는 것을 알 수 있었다.



첫 번째 연구에서 간접적인 언급 또한 많았다는 것을 알 수 있었기 때문에 두번째 연구에서 사용한 키워드만으로는 정확하게 검증할 수는 없었다. 따라서 직접적인 키워드 언급이 많은 달이라면 간접적인 언급이나 해시태그 언급과 같은 것들 또한 많다고 가정을 하여 연구를 진행하였다. 따라서 두 번째 연구에서는 도지코인 그래프가 상승한 달과 일론 머스크의 키워드 언급 트윗이 밀집한 달이 대부분 일치하는 모습을 보였다는 것을 알 수 있었다. 물론 이 그래프들만으로는 둘의 상관관계를 증명할 수 없다. 하지만 미국 경제전문매체인 마켓인사이더에 따르면 여론조사기관 피플세이가 2021년 2월 6일 ~ 2021년 2월 8일에 3만400명의 미국인의 상대로 머스크에 대한 의견을 묻는 여론조사를 실행한 결과 응답자의 37%가 일론 머스크의 트윗을 보고 한 번 이상을 투자했으며, 트윗을 기반으로 여러 번 투자한 응답자는 전체의 16%였다고 한다. 비록 여론조사이지만 두 그래프와 함께 비교해본다면 일론 머스크의 트윗과 도지코인의 상관관계에 대한 신뢰성을 높일 수 있다고 판단하였다.

결론적으로 간접적인 언급이나 해시태그, 이모티콘 사용 등으로 키워드를 직접적으로 언급한 트윗을 가지고만 판단했다는 것에서 연구의 한계점은 있었지만, 앞서 언급한 가정을 바탕으로 연구를 분석한 결과 일론 머스크의 개인 SNS인 트위터 트윗 내용이 도지코인과 같은 가상화폐의 가격에 상당한 영향을 주었음을 알 수 있었다.

**Ⅵ. 참고문헌**

- 황원지, “미국인 37%, 일론 머스크의 트윗 보고 투자했다”, Chosun Biz, 2021-03-11

- 이기우, 어나니머스, 머스크에 경고 “당신 장난에 많은 인생 파괴” , 조선일보, 2021-06-06

- [Python] snscrape 패키지를 이용하여 id / 기간별 특정 키워드가 포함된 트위터 자료 수집 , <https://shlee1990.tistory.com/1077>

- Python2.net , <https://www.python2.net/>

- stackoverlfow, <https://stackoverflow.com/>

- DelftStack, <https://www.delftstack.com/ko/>

- 왕초보를 위한 Python: 쉽게 풀어 쓴 기초 문법과 실습, <https://wikidocs.net/book/2>