

이미지 전이 학습과 생성 알고리즘을 활용한 NFT 작품 가격 예측

조이안¹, 김지윤¹, 한찬희², 김현희^{1,*}

¹ 동덕여자대학교 정보통계학과

² 동덕여자대학교 컴퓨터학과

20181194@dongduk.ac.kr, jytt7173@gmail.com, cksqml4631@gmail.com, heekim@dongduk.ac.kr

NFT Price Prediction Using Image Transfer Learning and Generative Adversarial Network

Leean Jo¹, Jiyeon Kim¹, Chanhee Han², Hyon Hee Kim^{*}

¹Dept. of Information and Statistics, Dongduk Women's University

²Dept. of Computer Science, Dongduk Women's University

요 약

대체 불가능한 토큰을 의미하는 NFT는 블록 체인 기반 기술 중 하나로 소유권과 거래 이력이 블록 체인에 기록이 된다는 장점을 지녀 미술 작품 거래에도 활발히 활용되고 있다. 하지만 현재 NFT 작품 거래 시장은 주식거래와 같이 회전율과 거래량을 중점으로 투기성이 짙으며, 작품의 섬네일 이미지만으로 거래 의사를 판단하는 경우가 많다는 문제점이 대두되고 있다. 따라서 작품의 다양한 특성에 대한 고려가 함께 이루어질 필요가 있다. 본 논문은 작품의 다양한 변수를 수집하여 최적의 변수 조합을 찾아내고, 이미지 전이 학습과 생성 알고리즘을 활용해 다양한 변수가 가격에 미치는 영향을 알아보고자 했다. 연구 결과 이미지 만을 활용한 가격 예측은 정확도가 높지 않다는 사실을 확인하였고 작품이 가진 여러 특성을 변수로 한 정형 데이터의 가격 예측 정확도가 더 높은 것을 알 수 있었다. 또한 생성 알고리즘을 통해 새로운 특성 조합의 작품들을 만들었고 가격을 수치 예측해보았다. 이를 실제 작가의 작품의 가격과 가격의 변동 추이를 대조해 작가의 작품 판매 가격 평균치 이상인 작품의 특성을 확인할 수 있었다.

1. 서론

NFT는 대체 불가능한 토큰을 뜻하는 블록 체인 기반 기술이다. NFT라는 새로운 유통 방식은 기술력, 수익창출의 확장은 물론, 시각 예술의 지평도 넓혔다. [1] 최근 세간의 큰 화제가 되고 있는 NFT는 섬네일 이미지만을 보고 작품을 구매하는 경향이 크기 때문에 사행성 오락과 다름이 없다는 비판을 받고 있다. NFT 투자를 높은 회전율, 거래량을 중점으로 주식 투기와 같은 접근하는 사람들이 많은 것이 현실이며 작품의 특성에 대한 고려는 충분하지 않은 실정이다[2]. 따라서 작품의 섬네일 이미지 외에도 NFT 작품의 가격과 투자 의사를 결정하는 다른 요소는 무엇인지에 관한 구체적인 연구는 부족하다. 물론 작품 가격과 거래량을 데이터로 활용하는 시계열 연구와, [3] 감성 분석을 활용한 멀티모달 연구가 있으나, [4] 작가의 작품이 지닌 각 특성들을 세세하게 분석한 연구는 이루어지지 않고 있다.

본 연구는 섬네일 이미지뿐만 아니라 NFT 거래 사이트에서 제공하는 다양한 특성을 변수로 활용하여 NFT 작품의 가격을 예측하였다. 회귀 예측과 변수 선택법을 활용하여 가격을 예측하기에 가장 적절한 변수 조합을 찾았으며 이미지 전이 학습을 통해 적은 이미지 데이터 수로 예측 성능을 높이고자 하였다. 작품의 가격을 NFT 거래 단위 중 하나인 이더리움의 4를 기준으로 작품 가격을 이진 분류하여 가격이 4 이더리움 이상인 작품과 그렇지 않은 작품을 예측하였다. 또한 생성 알고리즘을 통해 작가의 특성을 활용한 새로운 이미지를 생성 해보았다. 이러한 이미지들을 가격 예측 알고리즘의 입력으로 넣어 수치 예측 결과를 얻었으며 수치 예측 값을 실제 작가의 유사한 작품의 가격과 비교해보았다.

실험 결과로 NFT 작품의 이미지만을 활용한 예측 정확도는 낮으며 따라서 정형 데이터로 예측 하는 것이 더 높은 정확도를 보였다. 따라서 이미지만을

가지고 구매를 하는 것이 적절한 투자 방법이 되기는 어렵다는 것을 보여준다. 또한 생성으로 조합된 이미지의 가격 예측을 통해 작품 속 특성들이 가격에 영향을 줄 수 있다는 것을 알 수 있었다.

2. 실험 설계

2022년 4월 작품의 가격을 기준으로 총 37개의 데이터를 수집하였다. 작가 자체의 특성도 가격 결정에 영향을 미치지만 데이터 수집에 어려움이 있었다. 또한 여러 작가를 한꺼번에 고려하기에 한계가 존재해 크립토 아트 특성이 뚜렷하게 나타나는 한 명의 작가(Gremplin)로 연구 범위를 좁혔다. 최근 거래량이 증가하고있는 NFT 주요 거래 플랫폼인 Opensea[5]에서 등록되어 있는 많은 작가를 선택적으로 수집할 경우 한 작가가 만든 작품의 가격들이 다양한 이유를 포괄해 설명하기 힘들다고 판단하였기 때문이다. NFT 작품의 속성은 제각각이라서 모든 변수를 통일하기가 쉽지 않다. 따라서, NFT 전문 뉴스 사이트에서 선정한 새롭게 뜨는 NFT 작가 리스트에서 작품 수가 100개 이상의 충분한 작가인 Gremplin 을 선정했다. 아래의 <표 1>는 데이터 출처, 수집한 작가, 수집한 작품 수, 수집 변수, 전처리 과정, 실험 환경을 요약하였다.

데이터 출처	Opensea
수집한 작가	Gremplin
작품 수	377 (2022년 4월 근래 거래 결정)
수집 변수	작품의 이미지, 작품 거래 횟수(sales), 작품의 좋아요 수(likes), 소유주 변경 횟수(transfer), 작품의 모든 특성(property : background, body, accessory1, accessory2, head, eyes, mouth, clothes, custom, name)
전처리 유무	likes, sales, transfer : 수치형 유지, 텍스트 형태인 property : 범주형 변환
실험 환경	Colab, GPU, R

<표 1> 실험 설계 요약.

3. 예측 모델 및 실험 결과

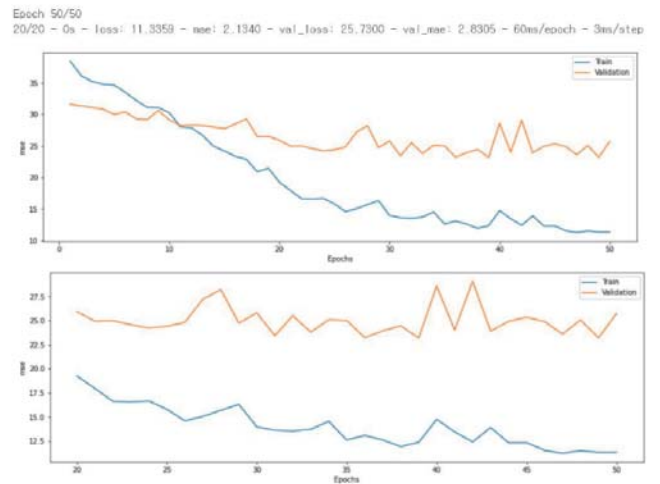
3.1 정형 데이터 변수 선택

작가의 작품을 판매하는 Opensea 사이트에서 데이터를 수집하고 각 항목별로 분류하여 변수를 상정하였다. (그림 1)과 같이 변수 선택법 중 단계 선택법을 적용하여 가격 예측에 가장 적절한 변수 조합을 확인하였다. 단계 선택법에는 제거된

학습변수들도 같이 나와서 선택된 변수들과 (-기호로 표현) 제거된 변수들을 (+기호로 표현) 한눈에 볼 수 있다.

price ~ sales + body_Val + transfers + head_Val + likes	
+	likes, head_Val, body_Val, transfers, sales
—	background_Val, accessory2_Val, eyes_Val, mouth_Val, accessory1_Val

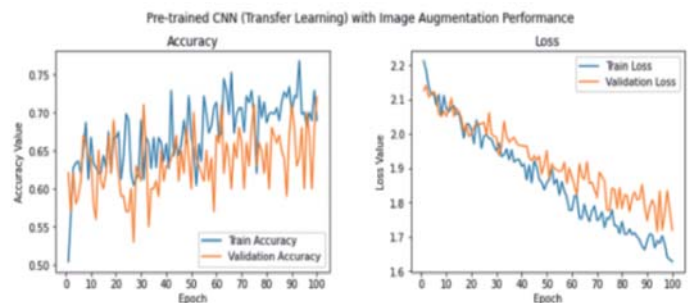
<표 2> 변수 선택법을 위한 단계 선택법.



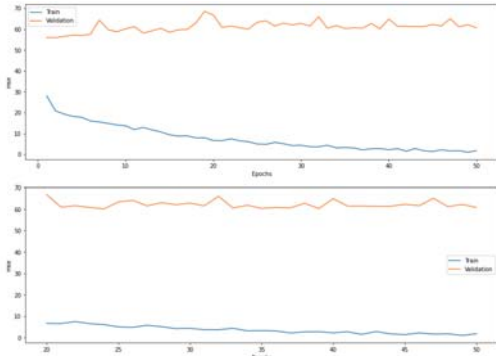
(그림 2) 변수 선택을 위한 회귀 예측.

따라서 최종적으로 선택된 변수 조합은 sales, body_Val, transfer, head_Val, likes 이다. 또한 가격 예측에 가장 영향을 주는 변수를 찾기 위해 회귀 예측을 실시하였다. (그림 2)는 MSE 값이 가장 작아지는 변수 조합을 위한 회귀 예측 결과를 나타낸다.

3.2 이미지 전이 학습을 통한 가격 이진 분류



(그림 3) VGG16 모델로 CNN 전이 학습 후 L2 규제.

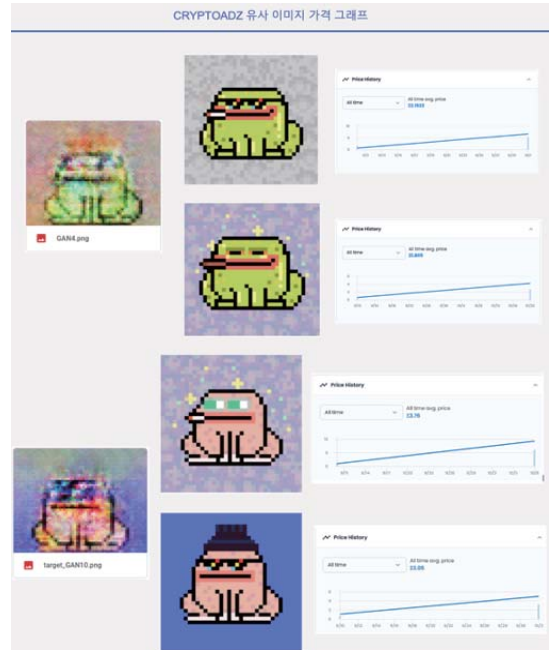


(그림 4) CNN 전이 학습 모델의 회귀 예측.

NFT 작품을 4 이더리움(NFT 거래 단위) 기준으로 이진 분류 예측하기 위해 전이 학습을 이용하여 CNN 모델을 생성했다. 적은 학습 이미지 데이터를 보완하기 위해 이미지 증식 통해 학습 이미지 한 개당 이미지 5 개를 생성하였고 사전 학습된 모델 중 VGG16 모델을 사용해 CNN 을 전이 학습했다. (그림 3) 그래프는 학습 이미지 생성 후 전이 학습 시킨 CNN 모델에 L2 규제를 적용한 그래프이며 학습된 모델의 가격 이진 분류 예측 정확도는 약 70%이다. 또한 (그림 4) 로 표현된 생성 이미지 가격 수치 예측의 경우, CNN 모델을 활용한 회귀 예측으로 출력층 이전의 모델의 구조는 앞서 언급한 모델과 동일하다.

3.3 생성 알고리즘을 활용한 작품 가격 예측

생성 이미지를 만들기 위한 입력 데이터 세트를 두 가지 그룹으로 나누었다. 4 이더리움 이상의 작품만을 활용한 생성 이미지 그룹과 모든 가격의 이미지를 포함한 생성 이미지 그룹이다. 이처럼 생성된 이미지를 가격 예측 알고리즘에 입력으로 넣었을 시 예측 값을 출력해보았다. 상위 가격이 나온 이미지와 유사한 작가의 작품과의 가격을 대조하였다. (그림 5)의 왼편에 있는 두 생성 이미지에 대응하는 작품을 각 2 개씩 비교하고 가격 이력을 확인해보았다. 상단의 두 작품 Cryptoadz#1738, Cryptoadz#2766 의 가격은 각각 6.669, 4.2 이더리움이다. 또한 (그림 5)의 하단의 두 작품인 Cryptoadz #6579, Cryptoadz#6888 의 가격은 각각 9.28, 5 이더리움이다. 그림 오른편의 작품 판매가 이력 그래프는 보통 우상향을 나타냄을 알 수 있었다. 이들은 작품 판매가 평균인 4 이더리움을 웃돈다.



(그림 5) 생성 이미지의 가격 예측 결과.

4. 결론 및 향후 연구

본 논문은 NFT 작품 가격 예측을 위한 변수를 변수 선택법, 회귀 분석 예측을 사용하여 탐색하였고, VGG16 모델을 활용한 이미지 전이 학습과 GAN 생성 알고리즘을 활용해 가격 예측을 진행했다. 실험 결과가 이 작가의 작품에서 해당 작가의 평균 판매가 이상을 지니는 작품들의 가격에 영향을 미치는 변수는 작품거래횟수, body_Val(작품 특성 중 몸의 형태와 색깔에 해당하는 변수), 소유주 변경 횟수, head_Val(작품의 머리 형태에 관한 변수), 작품의 좋아요 수로 나타났다. 또한 이미지만을 활용한 가격 예측 보다 다양한 특성 변수를 활용한 가격 예측 성능이 좋은 것으로 보아 작품을 구매할 때 다른 요소들도 고려하는 것이 좋을 것으로 생각된다. 향후 연구에서는 추가적으로 작가와 관련된 변수도 추가할 방안을 고려해야하며, 작품 가격에 영향을 주는 변수를 찾는 변수 선택법을 한 작가의 작품들에서 opensea 의 모든 작품으로 확대할 예정이다.

참고문헌

- [1] 권은용, “NFT 아트와 미술시장 유통의 변화에 대한 전망”, 성균관대학교 트랜스미디어 연구소, 제 1 권, 13 호, 1-16, 2022
- [2] 성소라 외, NFT 레볼루션, 더퀘스트, 2021
- [3] Bouabdallah A., “Multimodal Approach for Cryptocurrency Price Prediction”, University of Koblenz, 2022
- [4] 장혜림 외, “딥러닝을 이용한 미디어 콘텐츠 NFT Collection 가격 예측”, 한국통신학회 동계종합학술발표회, 평창, 2022, 1571-1572
- [5] 권혁준 외, “NFT의 거래 가능성 및 확장성에 대한 고찰”, 지급결제학회지, 제 13 권, 1 호, 257-272, 2021