

LSTM 모델을 이용한 조각투자 상품의 가격 예측: 뮤직카우를 중심으로

정현조

서울과학기술대학교 경영대학
(mrjung0987@seoultech.ac.kr)

이재환

서울대학교 경영연구소
(jaehwan@snu.ac.kr)

서지혜

서울과학기술대학교 경영대학
(jihae@seoultech.ac.kr)

최근 고액의 실물자산이나 채권을 분할하여 여러 투자자가 공동으로 투자하는 이른바 조각투자가 인기를 얻고 있다. 2016년 설립된 뮤직카우는 음원 유통에 따른 저작권료 참여 청구권을 조각투자할 수 있는 서비스를 세계 최초로 시작하였다. 본 연구에서는 딥러닝 알고리즘 중 하나인 LSTM 모델을 사용하여 뮤직카우에서 거래되는 저작권료 참여 청구권의 가격을 예측하는 연구를 진행하였다. 청구권의 이전 가격과 거래량, 저작권료와 같은 청구권과 관련된 변수 외에도, 음악 저작권료 참여 청구권 시장 상황을 나타내는 종합 지표와 경제 상황을 반영하는 환율, 국고채 금리, 한국종합주가지수도 변수로 사용하였다. 연구 결과 상대적으로 거래량이 낮은 조각투자의 사례에서도 LSTM 모델이 거래가격을 잘 예측하는 것을 확인할 수 있었다.

주제어 : LSTM, 조각투자, 거래가격 예측, 뮤직카우, 예측 모델

논문접수일 : 2022년 9월 2일

논문수정일 : 2022년 10월 4일

게재확정일 : 2022년 10월 31일

원고유형 : Regular Track

교신저자 : 서지혜

1. 서론

최근 MZ세대를 중심으로 미술품, 부동산 등을 분할하여 소액으로 투자할 수 있는 ‘조각투자’가 화제이다(주형연, 2022). 조각투자는 여러 명의 투자자가 실물자산이나 재산적 가치가 있는 권리를 분할하여 그 지분이나 청구권을 거래하는 새로운 투자방법이다(이영경, 2022). 국내 조각투자 시장은 2017년 소개된 이후 꾸준히 성장하여 2022년 초 기준 누적 공모 금액은 약 5천억 원 규모에 이른다(하나금융경영연구소, 2022). 국내 조각투자 플랫폼 중 하나인 ‘뮤직카우¹⁾’는

세계 최초로 음악 저작권에서 발생 되는 수익을 받을 권리인 저작권료 참여 청구권의 거래를 중개한다. 뮤직카우는 누적 회원 수 100만 명을 돌파했고 누적 거래액은 3,399억에 이를 정도로 큰 인기를 끌고 있다(이경미, 2022). 뮤직카우는 자회사를 통해 저작권의 일부를 구매 한 뒤, 음악 저작권에서 발생 되는 수익을 받을 권리인 저작권료 참여청구권을 양도받아 쪼개 경매를 통해 투자자들에게 판매하고 그것을 뮤직카우 내의 마켓에서 자유롭게 매수, 매도할 수 있도록 한다(허란, 2022). 투자자들은 뮤직카우를 통해 분배되는, 저작권료 참여청구권의 저작권료에서 받

1) <https://www.musicow.com/>

* 본 연구는 서울과학기술대학교 교내연구비지원에 의해 이루어졌습니다

생하는 수익이나, 매매 간 차익을 통해 수익을 얻을 수 있다.

현재까지 주가 예측이나 비트코인과 같은 암호화폐의 가격 예측은 많은 연구가 진행되어 온 반면 (김선웅, 2021a, 김선웅, 2021b; 이준식 등, 2018), 신생분야라고 할 수 있는 조각투자자에 대한 연구는 드물다. 따라서 본 연구는 저작권료 참여 청구권을 거래하는 플랫폼인 뮤직카우를 대상으로 저작권료 참여청구권 가격 예측에 대한 연구를 수행하고 이를 통해 조각투자자에 대한 이해를 높이고자 한다. 가격 예측에는 딥러닝 알고리즘 중 LSTM(Long Short-Term Memory) 알고리즘을 사용하였고, 조각투자 거래와 관련된 변수, 조각투자 거래 시장과 관련된 변수, 거시경제와 관련된 변수가 예측 모델에 포함되었다. 이 논문은 다음과 같이 구성되어 있다. 2장에서는 배경 연구로서 조각투자자와 관련된 기존 연구와 투자 상품에 대한 거래가격을 예측하는 연구들을 살펴본다. 3장에서는 본 연구에서 사용한 데이터에 대해서 자세히 알아보고 LSTM에 기반한 뮤직카우 가격 예측 모델을 소개한다. 4장에서는 제안한 예측 모델의 결과를 살펴보고, 마지막 5장에서 본 연구가 시사하는 시사점 및 한계를 논의한다.

2. 배경연구

2.1 조각투자

조각 투자의 정의는 하나의 자산에 다수의 투자자들이 다같이 투자하여 이익을 공동으로 배분받는 투자 방법이다(김인경, 2022; 정재욱, 2022). 이런 조각투자의 대표적 대상으로는 미술

품이 있다. 아트 바젤과 UBS에서 펴낸 ‘아트마켓 2021’ 보고서에 의하면, 2019년 대비 2020년 전 세계 미술 시장의 규모는 22% 감소했으나 온라인을 통한 거래액은 2배 이상 성장해 사상 최고액인 124억 달러(한화 약 14조6,000억 원)를 기록한 것으로 나타났다(Art Basel, 2021). 우리나라도 이러한 경향이 나타났는데, 아트와 재테크를 합친 아트테크 즉, 온라인을 통해 미술품에 투자하는 것에 관심이 집중되고 있고 특별히, 온라인 환경에 익숙한 MZ세대가 이러한 미술시장에 많이 유입되고 있다(김인경, 2022; A Team, 2021, Kim, 2021). 가장 대표적으로 떠오르고 있는 방식이 ‘미술품 분할소유권,’ 또는 ‘미술품 공동구매 투자’이다. 이는 적게는 수천만 원 많게는 수십억 원을 호가하는 미술품의 지분을 여러 조각으로 나눠 주식처럼 투자하는 방식을 일컫는다. 투자 기본 금액은 플랫폼마다 각각 다르지만 적게는 1,000원에서 10만 원 등 비교적 소액으로 시작할 수 있고 (강서구, 2021; A Team, 2021; Kim, 2021) 이러한 적은 금액으로 앤디 워홀, 김환기 등 국내외 유명 근현대 작가들이 작품을 소유할 수 있다(강서구, 2021; 이익진, 2021). 이러한 아트테크 플랫폼들이 크게 성장하는 이유는 다음과 같다. 먼저, 코로나19의 여파로 시장이 빠른 속도로 온라인으로 옮겨간 것이 크고 다음으로 고(故) 이견희의 미술품 국가 기증이나 케이팝(K-Pop) 그룹 BTS의 멤버 RM의 미술관 기증 등 유명인들의 미술품에 대한 관심이 회자되면서 MZ 세대들의 관심도 급 승했다는 분석이 있다. 그러나 경제적인 면을 고려해 보면, 저금리 기조로 인해 국내 금융시장이 불안정해졌고 그 때문에 젊은 세대들이 비교적 안정적인 투자처를 찾고 있는 것을 가장 큰 이유로 보고 있다(강서구, 2021).

또 다른 조각투자 대상은 뮤직카우로 대표되는 음악 저작권료 청구권이 있다. 음악 저작권자는 음악 저작물 사용료 징수 규정에 따라 저작권료를 지급받을 수 있다. 국내 음악 저작물 사용료 징수 규정은 한국음악권저작협회, 함께하는 음악저작인협회, 한국음반산업협회, 한국음악실연자연협회 등 4개 신탁 관리 단체가 협의해 만들었으며 문화체육관광부의 승인을 받은 후 시행된다. 규정에 따르면 음원으로 발생한 수익은 서비스사업자, 음반제작자, 저작권자, 실연자에게 분배된다. 음원의 판매 형태(예. 다운로드, 스트리밍)에 따라 비율이 다르지만 평균적으로 저작권자는 음원유통 수익의 10% 가량을 분배받는다. 뮤직카우는 음원유통에 따라 저작권자에게 분배되는 수익에 대한 권리를 거래하는 플랫폼이다. 투자자는 권리에 대한 거래를 통해 차익을 얻거나, 음원의 인기에 따른 저작권료 지급의 두 가지 방식으로 수익을 올릴 수 있다. 주식시장에 비교하자면 저작권료는 일종의 배당금으로 생각할 수 있다.

음악 저작권료 청구권 투자는 원저작자가 자신의 곡 저작권 일부를 공개하면 대중들이 경매로 자유롭게 저작권료 지분을 구매하는 방식이다. 대상은 현재 활동하는 유명 가수들의 노래로 2021년 6월 거래된 음원만 1,200여 곡 정도이다(김인경, 2022; 이형민, 2022). 2018년 1월부터 2021년 6월까지 뮤직카우에서 거래된 음악 저작권료 참여 청구권의 연간 배당 수익률은 평균 6.87%이며 음원 수익을 나눠 가질 수 있다는 매력 및 안정적이며 높은 수익률에 대한 기대로 2022년 4월 기준 누적 회원수는 110만명, 누적 거래 규모는 3715억원을 돌파했다(김인경, 2022). 예를 들어 보면, 2022년 4월 진행된 경매에서 아이유의 ‘봄 안녕 봄’ 저작권료 참여 청구권 1조각은

평균 61,600원에 낙찰되었고 6월 기준 78,000원에 거래 중이다. 모든 음원이 다 좋은 수익성을 보이는 것은 아니지만 평균적으로 음원은 발매된 해에 가장 많은 저작권료가 발생하고 대중성을 확보한 음원의 경우 발매 후 일정 수준 저작권료가 감소하다가 2년 후 점차 안정적인 수익을 내는 구조가 된다고 한다(김인경, 2022).

조각투자자와 관련된 연구는 국내의 경우 법이나 규제 측면에서 이루어진 연구(김인경, 2022; 이영경, 2022; 정재욱, 2022)가 주로 존재하며 해외의 경우 부동산 조각투자자와 관련된 연구(Lowies et al., 2018)가 주로 존재한다.

2.2 거래가격예측

머신러닝(machine learning)을 통해 금융시장의 비선형 구조를 식별할 수 있다는 연구(Huck, 2009; Huck, 2010)가 발표된 후, 많은 연구들이 주가를 예측하고자 SVM(Support Vector Machine), MLP(Multi-Layered Perception), DNN(Deep Neural Network), RNN(Recurrent Neural Network), LSTM(Long Short-Term Memory) 등을 사용해왔다(Huck, 2009; Huck, 2010). Devadoss & Ligori (2013)는 MLP와 ANN(Artificial Neural Network)과 역전파 알고리즘을 이용하여 봄베이 증권거래소에서 선택한 섹터의 주가를 예측하는 연구를 하였다. 허준영과 양진용(2015)은 SVM을 이용하여 기업의 재무정보를 기반으로 주가를 예측하였다.

신동하 등(2017)은 DNN과 RNN, LSTM을 사용하여, 주가의 시가, 고가 등을 이용한 기술 분석 데이터와 환율, 전산업생산지수 등의 거시경제지표를 이용하여 주가 예측률 향상을 위한 딥러닝 모델을 연구하였다. 이는 과거 학습 결과를 현재 학습에 사용하는 딥러닝 네트워크인 RNN

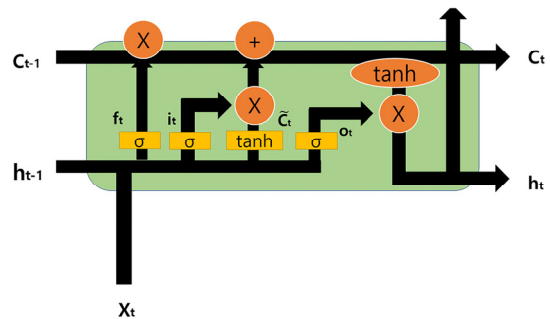
과 RNN의 과거 학습결과를 Cell state를 통해 전달하여 기울기 소실을 해결한 LSTM이, DNN에 비해 빠르고 안정적으로 학습하였고, 예측을 항상 보이는 것을 확인하였다. 이후 주일택과 최승호(2018)는 주가와 거래량 데이터로 양방향 LSTM을 이용하여 주가를 예측함으로써 단방향 신경망일 때보다 오차를 개선하였다. 한태동(2021)은 LSTM을 이용하여 주가를 예측할 때, 주가에 영향을 미치는 내부 요소, 외부요소 모두를 고려하기 위해 기술적 지표, 거시 경제 지표, 시장 심리 지표를 활용하였다.

암호화폐 중에서 가장 시가총액이 높은 비트코인 또한 딥러닝 기술을 통해 가격을 예측하는 선행 연구들이 진행되어 왔다. McNally et al. (2018)은 비트코인 가격만을 이용하여 베이지안 최적화 RNN과 LSTM을 통해 비트코인 가격을 예측하였다. Aggarwal et al. (2019)는 금 가격과 트위터 감성분석을 매개변수로 사용하여 CNN, LSTM, 및 GRU(Gated Recurrent Unit)를 이용해서 비트코인 가격을 예측했다. 안유진과 오하영(2021)은 LSTM을 기반으로 on-chain 데이터를 활용해서 비트코인의 가격을 예측하였다.

시계열 데이터 분석에 활용되는 대표적인 딥러닝 모델로는 RNN(Recurrent Neural Network)이 있다. RNN 모델에서 출력 값은 다음 단계 연산을 위한 입력 값으로 반복적으로 사용되는데, 은닉층을 이용해 과거의 정보를 저장하고 이용한다는 특징을 지닌다. RNN은 활성화함수로 하나의 하이퍼볼릭 탄젠트(Tanh)의 활성화 함수를 주로 사용하는데, 이로 인해 타임 스텝이 길어지면 과거의 학습 결과가 사라지는 장기 의존 문제를 가지고 있다. 이를 개선한 대표적인 모델로 LSTM(Long Short-Term Memory)이 있다.

LSTM 모델은 내부에 셀(Cell)의 집합을 가지

고 있다. 각 셀은 데이터의 순서를 기억하는 특성이 있다. 이전의 셀 스테이트(Cell State)인 C_{t-1} 가 큰 변화 없이 다음셀 스테이트인 C_t 로 전달되기 때문에 장기 의존 문제를 개선한다. 이러한 셀은 게이트(gate)로 통제하는데 게이트는 Forget gate, Input gate, Output gate 등이 있다. <그림 1>을 보면 Forget gate(f_t)는 0과 1사이의 값을 출력하는데 1인 경우 기억하고 0인 경우 잊어버린다. Input gate(i_t)는 셀에 저장할 정보를 선택한다. Output gate(o_t)는 셀 스테이트 값과 다른 정보들을 기반으로 각 셀에서 생성되는 값을 결정한다.



<그림 1> LSTM의 구조

3. 연구의 설계

3.1 변수 선정

선행 연구에 기술한 것과 같이, 주식의 예측을 위해 활용된 변수들은 주가와 거래량 등의 기술 분석 데이터, 환율, 전산업생산지수, 금리, KOSPI 등의 거시경제 지표 그리고 시장심리지표 등이었다(신동하 등, 2017; 안유진, 오하영, 2021; 주일택, 최승호, 2018; 한태동, 2021; Aggarwal et

al., 2019; McNally et al., 2018). 비트코인 예측을 위해 활용된 변수도 비트코인의 가격과 금 가격 트위터 감성분석 데이터, on-chain 데이터 등이 사용되었다(Aggarwal et al., 2019). 선행연구에서 활용했던 변수들을 <표 1>에 정리하였다.

<표 1> 선행연구에서 사용된 변수

연구	예측	변수
신동하 등 (2017)	주가	시가, 종가, 고가, 저가, 거래량, 환율, 산업생산지수 Commodity Channel Index
주일택, 최승호 (2018)	주가	시가, 종가, 고가, 저가, 거래량
한태동 (2021)	주가	거래량, 코스피 지수, 환율, 이자율, 특정 종목에 대한 인터넷에서의 검색 빈도
안유진, 오하영 (2021)	비트코인 가격	Exchange Reserve, Exchange Transactions Count Outflow, Addresses Count Inflow, Fund Flow Ratio, Estimated Leverage Ratio, Stablecoin Supply Ratio, Miner's Reserve, Miner's Reserve in USD, Open Interest, Hashrate
Aggarwal et al. (2019)	비트코인 가격	비트코인 가격(5분 단위), 금 가격(일 단위), 트위터
McNally et al. (2018)	비트코인 가격	시가, 종가, 고가, 저가, 비트코인 채굴 난이도

이러한 선행연구에서 활용한 변수들을 참고하여 대부분의 연구에서 공통적으로 활용한 과거 가격과 기술지표 2가지, 거시경제 지표 3가지를 선택하였고 (안유진, 오하영, 2021; Aggarwal et al., 2019; McNally et al., 2018) 여기에 음악 저작권료 청구권만이 가진 저작권료도 변수로 선정하였다.

모델에 사용된 변수에 대한 설명은 <표 2>와 같다. 가격 변수로는 음악 저작권료 청구권의 증가를 사용하였다. MCPI는 2019년 1월 1일의 음악 저작권료 청구권 평균 가격을 기준으로 만들어진 지수로 2019년 1월 1일의 MCPI는 100이다. 2022년 1월 1일의 MCPI는 250.02인데, 이는 2년 간 음악 저작권료 청구권의 평균 가격이 150% 증가했다고 해석할 수 있다. 음악저작권료는 음원 유통으로 발생하는 수익에서 음악 저작권료 청구권을 가진 투자자에게 배분되는 수익을 의미한다.

<표 2> 변수 설명

변수	변수 설명
가격(P)	종가
거래량(V)	거래량
MCPI(M)	Music Copyright Property Index 지수
음악저작권료(C)	음원 유통으로 발생하는 수익
환율(E)	원/미국달러 환율(매매기준율)
금리(I)	3년만기 국고채 금리
KOSPI(K)	한국종합주가지수

3.2 연구 모델

본 연구의 목적은 위에서 선정된 변수들을 이용하여 개별 저작권료 청구권의 가격을 예측하는 것이다. 예측하고자 하는 가격의 시점은 $t+1$ 일부터 $t+5$ 일까지의 가격($\{P_{t+1}, \dots, P_{t+5}\}$)으로 하였다. 예측하기 위해 사용되는 변수들의 시점은 30일 단위인 t 일부터 $t-29$ 일까지로 하였다. 이를 수식으로 나타내면 <식 1>과 같다.

$$\begin{aligned}
 & \begin{Bmatrix} P_{t+1,i} \\ \dots \\ P_{t+5,i} \end{Bmatrix} \\
 & = f \left[b + w \times \begin{Bmatrix} P_{t,i} & P_{t-1,i} & \dots & P_{t-28,i} & P_{t-29,i} \\ V_{t,i} & V_{t-1,i} & \dots & V_{t-28,i} & V_{t-29,i} \\ C_{t,i} & C_{t-1,i} & \dots & C_{t-28,i} & C_{t-29,i} \\ M_t & M_{t-1} & \dots & M_{t-28} & M_{t-29} \\ E_t & E_{t-1} & \dots & E_{t-28} & E_{t-29} \\ I_t & I_{t-1} & \dots & I_{t-28} & I_{t-29} \\ K_t & K_{t-1} & \dots & K_{t-28} & K_{t-29} \end{Bmatrix} \right] \\
 & \dots \text{ <식 1>}
 \end{aligned}$$

<식 1>에서 $P_{t,i}$ 은 시점 t 에서 음악 i 의 음악 저작권료 청구권 가격을 나타내고, $V_{t,i}$ 는 시점 t 에서 음악 i 의 음악 저작권료 청구권의 거래량, $C_{t,i}$ 는 시점 t 에서 음악 i 의 저작권료, M_t 는 시점 t 의 MCPI 지수, E_t 는 시점 t 의 환율, I_t 는 시점 t 의 금리, K_t 는 시점 t 의 KOSPI 지수를 나타낸다.

본 연구에서는 LSTM을 사용하여 음악 저작권료 참여 청구권 가격 예측 연구를 실행하였다. 과거의 정보를 기억해서 미래의 입력정보로 넘겨주어 가격을 예측하는 특성을 가지고 있기 때문에, LSTM은 주가와 같은 장기간의 비선형 시계열을 예측하는데 좋은 성능을 보여주고 있다. 본 연구에서도 3년 정도의 긴 기간의 데이터를 바탕으로 음악 저작권료 참여 청구권을 예측하는 것이 목적이기 때문에 LSTM이 연구에 적합하다고 판단하였다.

3.3 데이터 수집 및 처리

음악 저작권료 청구권의 개별 가격 데이터와 기술지표인 거래량과 MCPI 지수, 그리고 월별 음악저작권료는 뮤직카우에서 수집하였다. 음악

저작권료 청구권의 개별 가격과 거래량, 월별 저작권료, MCPI 지수는 뮤직카우 이용자들 간의 거래가 이루어지는 뮤직카우 마켓²⁾에서 파이썬 (Python 3.8)과 *BeautifulSoup* 패키지와 *Selenium* 패키지를 이용해서 수집하였다. 그 후, 뮤직카우에서 분류한 저작권의 ID를 통하여 각각 구분하였다. 거시경제 지표인 환율과 금리는 한국은행 경제통계시스템에서 제공하는 환율 데이터와 3년 만기 국고채 금리를 이용하였고 KOSPI는 KOSIS 국가통계포털을 통해 수집하였다.

각각의 데이터들을 출처로부터 수집하고 전처리 하였다. MCPI 지수가 2019년 1월 1일을 기준으로 만들어졌기 때문에 학습 시킬 데이터의 기간은 2019년 1월 1일부터 2022년 2월 28일까지로 정하였다. 따라서 가격과 거래량 데이터는 2019년 1월 1일 전에 뮤직카우 옥션을 통해 상장된 123개 곡의 2019년 1월 1일부터 2022년 2월 28일까지의 거래 데이터를 사용하였다. 총 1,155일의 일별 데이터로 고가, 저가, 증가, 변화율로 수집되었으나 본 연구에서는 증가만 분석에 사용하였다. MCPI 지수도 2019년 1월 1일부터 2022년 2월 28일까지의 총 1,155일 데이터를 수집하였다.

거시경제지표인 환율과 금리, KOSPI는 주말과 공휴일에는 데이터 값이 없는 경우를 감안하여 2018년 12월 31일부터 2022년 2월 28일까지의 데이터를 수집하였고 비어 있는 값들은 직전 데이터 값으로 대체하였다. 그 후 기간을 2019년 1월 1일부터 2022년 2월 28일까지로 조정하였다. 음악 저작권 지표인 음악저작권 자료는 월별로 2019년 1월부터 2022년 2월까지의 데이터를 수집하고, 뮤직카우에서는 첫 영업일에 저작권료를 정산하므로, 월별 데이터를 일별데이터로 바꾸되, 달이

2) <https://www.musicow.com/market>

바뀔 때 저작권료가 바뀌는 게 아니라 첫 영업일을 기준으로 저작권료가 바뀌도록 하였다. 마지막으로 딥러닝에서는 입출력의 데이터의 크기가 작아야 학습이 효과적이기 때문에 <식 2>와 같이 대표적 정규화 기법인 최소-최대 정규화(Min-Max Normalization)을 이용하여 모든 입출력 데이터를 0과 1사이로 정규화 하였다(신동하 등, 2017).

$$f(x) = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \dots \text{<식 2>}$$

4. 실험 및 결과

4.1 검증 및 파라미터 튜닝

본 연구에서는 2019년 1월 1일부터 2022년 2월 28일까지 총 1,155일 기간의 데이터 중 30일을 하나의 인스턴스로 하고 5일의 데이터를 예측하는 것을 목적으로 했을 때 전체 인스턴스는 1,120개가 된다. 이러한 인스턴스를 비율은 8:2로 하여 896개의 훈련 데이터 셋과 224개의 테스트 데이터 셋으로 나누었다. 예측된 저작권료 청구권 가격과 실제 예측된 저작권료 청구권의 가격의 오차율을 확인하기 위하여 평균 제곱근 오차인 RMSE(Root Mean Square Error) <식 3>을 사용하였다. 수식에서 Y_i 는 예측 값, X_i 는 실제 값을 나타낸다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - X_i)^2} \dots \text{<식 3>}$$

LSTM의 하이퍼파라미터로는 Learning rate, mini-batch size, Drop-out 등이 있다(신성호 등, 2018). Learning rate는 예측 값과 실제 값의 차이

가 날 때 갱신하는 가중치의 크기를 의미하며, 지나치게 작은 경우 학습 시간이 오래 걸리고 지나치게 크면 발산할 수 있다. 본 연구에서는 Learning rate를 일반적으로 많이 사용되는 0.001로 설정하였다. mini-batch 방식은 한 번에 모든 데이터를 활용하는 full-batch와 달리, 전체 학습데이터 셋을 서브 데이터 셋으로 나누어 여러 번 학습하는 방식이다. mini-batch 사이즈는 일반적으로 50에서 100 정도의 크기가 많이 사용되므로, 본 연구에서는 50으로 정하였다. 따라서 하나의 학습데이터를 한 번 학습(Epoch) 할 때, 18번의 학습과정(iteration)이 필요하다. Drop-out은 학습 모델이 학습 데이터에 과적합(over-fitting) 되는 문제를 방지하기 위해 사용되는 기법으로 그 값을 0.01로 정하였다. 예측값과 실제값의 차인 cost function을 최소화하는 기법인 optimizer는 최근 많이 사용되는 adam optimizer를 사용했고, Epochs는 학습을 얼마나 반복하는지를 의미한다. 본 연구해서 실험해본 결과 epoch가 100회를 넘어가면 loss 값이 더이상 감소하지 않는 모습을 보였기 때문에 Epochs는 100으로 하였다. 본 연구의 LSTM에 사용된 하이퍼파라미터를 <표 3>에 정리하였다.

<표 3> 변수 설명

하이퍼파라미터	값
Learning rate	0.001
mini-batch size	50
Drop-out	0.01
Optimizer	Adam
Epochs	100

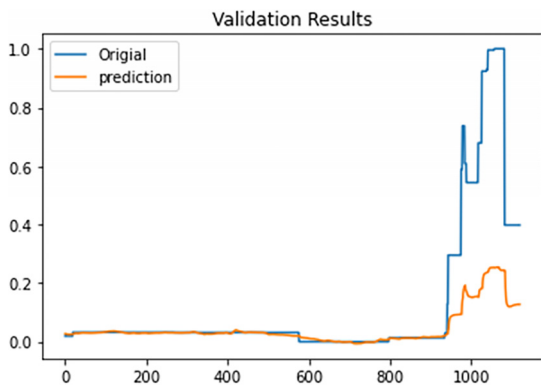
4.2 실험 결과

하이퍼파라미터 값을 바탕으로 실험을 하여 123개의 곡의 가격을 예측하고 RMSE를 이용하여

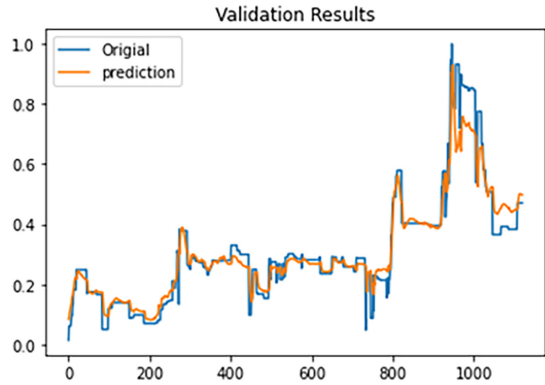
Train error와 Test error를 각각 계산하였다. 123개의 Test-error 중 최솟값은 0.04872이고 1분위 값은 0.08086이다. 중앙값은 0.09904이며 평균값은 0.11443이었다. 3분위 값은 0.135이고 최댓값은 0.44517이다. Test error에 대한 기초 통계량 값은 <표 4>에 정리하였다. 모델의 예측 성능을 더 쉽게 이해할 수 있도록 Test error의 값이 최댓값, 중앙값, 최솟값에 해당하는 음원의 실제 데이터와 예측 데이터를 시각화하여 각각 <그림 2>, <그림3>, <그림4>로 나타내었다. 각 그림의 가로축은 time을 나타내고 세로축은 정규화된 음악 저작권료 청구권 가격을 나타낸다. 또한 각 그림에서 파란색 선은 음악 저작권료 청구권의 실제 거래 가격을 의미하고 주황색 선은 예측값을 의미한다.

<표 4> 예측 오차의 기초 통계량

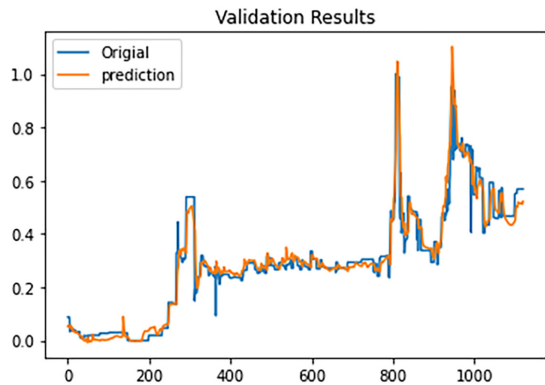
구분	Train error	Test error
최솟값	0.00480	0.04872
1분위수	0.01826	0.08086
중앙값	0.02627	0.099036
평균	0.02751	0.11443
3분위수	0.03444	0.13500
최댓값	0.09359	0.44517



<그림 2> 모델 성능(RMSE=0.44517, 최댓값)

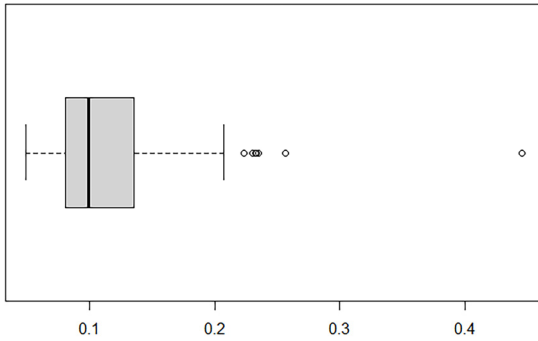


<그림 3> 모델 성능(RMSE=0.09904, 중앙값)



<그림 4> 모델 성능(RMSE=0.04872, 최솟값)

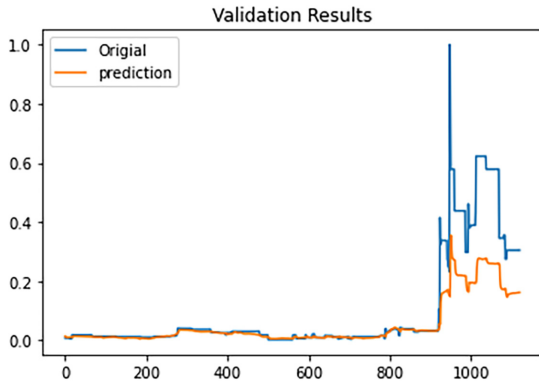
중앙값을 보면 비교적 예측 값이 가격을 추종하는 편이나, 최댓값을 포함한 몇몇 저작권료 청구권의 가격은 예측이 잘 되지 않았음을 볼 수 있다. 예측이 잘 되지 않은 값들을 확인하기 위해 <그림 5>와 같이 상자 그림을 그리고 이상치(outlier)를 분류한 결과, 6개의 곡이 이상치로 도출되었다. 상자 그림에서 가로축은 Test error를 나타낸다. 이상치에 해당하는 모델들의 성능을 그래프를 이용해 확인해보았다. <그림 6>과 <그림 7>은 6개의 이상치 중 2개 이상치에 대한 모델 성능을 나타낸 그래프이다.



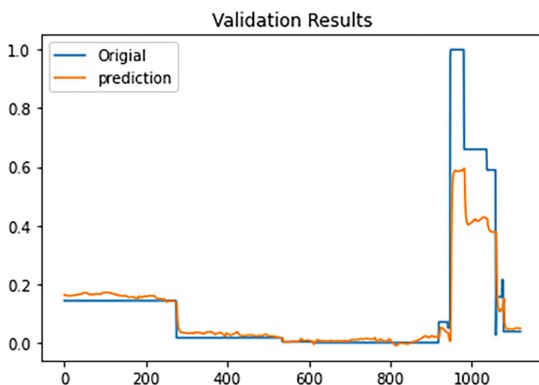
<그림 5> Test error 상자 그림

이상치에 대한 그래프를 살펴보면 특정 구간에서 급격한 가격 변동이 발생했을 때 예측 정확도가 떨어지는 것을 볼 수 있다. 수집한 123개의 저작권료 청구권의 일평균 거래량이 약 1.53회 정도로 매우 낮은 수치이기 때문에 이러한 급작스러운 가격 변동이 발생할 수 있고, 그로 인해 예측 정확도가 떨어진 것으로 추측할 수 있다.

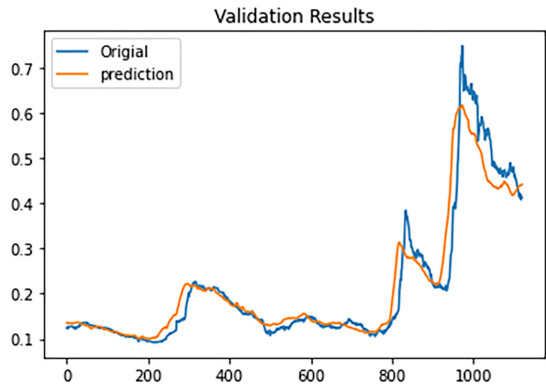
개별 저작권이 아닌, 123개의 저작권료 가격의 평균도 예측하여보았다. 개별 저작권료 가격, 개별 저작권 거래량, 개별 음악저작권료를 정규화한 123개의 저작권료 청구권의 평균치로 대체하였다. 그 외 MCPI지수나 환율, 3년만기 국고채 금리, 코스피 지수는 동일하게 사용하였으며, LSTM의 하이퍼파라미터 또한 유지하며 같은 조건에서 실험하였다. 그 결과 RMSE는 0.089447이 나왔고, 123개의 RMSE 평균보다, 123개 평균 가격의 RMSE가 더 낮다는 것을 알 수 있다. 아래 <그림 8>은 그래프로 시각화한 것이다. 주황색 선은 예측 데이터이고 파란색 선은 실제 데이터이다.



<그림 6> 모델 성능(RMSE=0.234861)



<그림 7> 모델 성능(RMSE=0.0.230345)



<그림 8> 123개 저작권 청구권의 평균 예측 모델

5. 결론

본 연구는 시계열 데이터의 분석에 적합한 머신러닝 기법 중 하나인 LSTM을 이용해 음악 저작권료 청구권의 거래 가격을 예측하는 모델을 개발하였다. 기존 연구들은 주로 주식과 가상화폐와 같은 거래량이 많고 대중적인 투자 수단의 가격을 예측하였다. 본 연구는 최근 많은 주목을 받고 있는 새로운 투자 방법인 조각투자 상품의 거래 가격을 예측했다. 비교적 많은 연구가 이루어지지 않은 새로운 분야의 연구라는 점에서 의의가 있다고 할 수 있다.

음악 저작권료 청구권은 비교적 새로운 투자 수단이며 단위 가격이 높아 조각투자 방식으로 투자가 진행된다는 특징이 있다. 국내에서 큰 인기를 끌고 있는 조각투자 플랫폼인 뮤직카우를 연구 대상으로 선택하여 음악 저작권료 청구권 가격을 예측하는 딥러닝 모델을 최초로 제시하였다. 2019년 1월 1일부터 2022년 2월 28일까지 총 38개월 간의 데이터가 실험에 사용되었고, 실험 결과 평균 일일 거래량이 많지 않다는 데이터의 특징에도 LSTM이 잘 작동한다는 것을 확인할 수 있었다.

본 연구의 한계점은 다음과 같다. 첫째, 음악 저작권료 청구권 가격 예측에 있어 음악 자체의 특성을 고려하지 못하였다. 음악 저작권료 청구권은 저작권료 수익을 배분받을 권리로, 음악 저작권료라는 내재적 가치를 지니고 있다. 음악 저작권료는 음원의 유통 즉, 인기에 비례하는데 음악의 인구나 대중성과 같은 요인이 변수에 포함되지 않았다. 향후 연구에서는 음악의 인기와 같이 음악 자체와 관련된 변수를 분석에 활용할 수 있을 것이다. 둘째, 본 연구에서는 2019년부터 거래가 시작된 123개의 곡을 이용해 실험을 하

였으나, 현재는 뮤직카우에서는 1,000개가 넘는 곡이 거래가능하다고 알려져있다. 따라서 분석 기간을 달리한다면 더 많은 데이터에 대하여 실험할 수 있을 것이다.

마지막으로 기존 전통적인 시계열 예측 기법과 비교 실험이 필요하다. 이를 통해 기존 방법보다 본 연구 방법인 LSTM의 성능이 더 높다는 것을 증명할 필요가 있을 것이다.

참고문헌(References)

[국내 문헌]

- 강서구. (2021, 10월 13일). [조각투자 명암⑥] “법적 사각지대서 미술품 굴리다”. 더스coop. <https://www.thescoop.co.kr/news/articleView.html?idxno=52178>
- 김선웅. (2021a). 딥러닝을 이용한 비트코인 투자 전략의 성과 분석. 한국융합학회논문지, 12(4), 249-258. <https://doi.org/10.15207/JKCS.2021.12.4.249>
- 김선웅. (2021b). 암호화폐 트레이딩시스템의 수익성 분석. 한국디지털콘텐츠학회 논문지, 22(3), 555-562.
- 김용찬, 박진수, 서지혜. (2018). An Algorithm for Finding a Relationship Between Entities: Semi-Automated Schema Integration Approach. 지능정보연구, 24(3), 243-262.
- 김인경. (2022). 조각투자 전성시대... 2030 러시. KISO 저널, (47), 34-37.
- 김재경, 서지혜, 안도현, 조윤호. (2002). 협업 필터링 기법을 활용한 개인화된 상품 추천 방법론 개발에 관한 연구 (A Personalized Recommendation Methodology based on Collaborative Filtering). 지능정보연구, 8(2), 139-157.

- 신동하, 최광호, 김창복. (2017). RNN과 LSTM을 이용한 주가 예측을 향상을 위한 딥러닝 모델. 한국정보기술학회논문지, 15(10), 9-16. <https://doi.org/10.14801/jkiit.2017.15.10.9>
- 신성호, 이미경, 송사광. (2018). LSTM 네트워크를 활용한 농산물 가격 예측 모델. 한국콘텐츠학회논문지, 18(11), 416-429. <https://doi.org/10.5392/JKCA.2018.18.11.416>
- 안유진, 오하영. (2021). On-Chain Data를 활용한 LSTM 기반 비트코인 가격 예측. 한국정보통신학회논문지, 25(10), 1287-1295. <https://doi.org/10.6109/jkiice.2021.25.10.1287>
- 이경미. (2022, 4월 4일). 뮤직카우 자본시장법 규제 가다...조각투자 지각변동. 한겨레신문. <https://www.hani.co.kr/arti/economy/finance/1037364.html>
- 이영경. (2022). 조각투자의 거래구조와 투자자 보호. 선진상사법률연구. (99), 37-75.
- 이익진. (2021, 10월 1일). 투자가 '예술'이네...아트테크를 아시나요. 비즈니스워치. <http://news.bizwatch.co.kr/article/market/2021/09/29/0030>
- 이준식, 김건우, 박도형. (2018). 비트코인 가격 변화에 관한 실증분석: 소비자, 산업, 그리고 거시변수를 중심으로. 지능정보연구, 24(2), 195-220.
- 정재욱. (2022). 조각투자와 뮤직카우 '제재 보류' 결정이 남긴 것들. 고시계, 67(6), 143-146.
- 주일택, 최승호. (2018). 양방향 LSTM 순환신경망 기반 주가예측모델. 한국정보전자통신기술학회논문지, 11(2), 204-208. <https://doi.org/10.17661/jkiict.2018.11.2.204>
- 주형연. (2022, 3월 31일). 조각투자에 몰린 MZ세대... '투자자 보호장치' 급선무. 세계비즈. <http://www.segyebiz.com/newsView/20220331515416>
- 하나금융경영연구소. (2022). 조각투자 이해하기. 하나Knowledge+ 제 14호
- 한태동. (2021). LSTM을 이용한 주가 예측: 기술 지표, 거시 경제 지표, 시장 심리의 조합을 중심으로. 융복합지식학회논문지, 9(4), 189-198. <https://doi.org/10.22716/SCKT.2021.9.4.055>
- 허란. (2022, 4월 5일). '증권성' 논란 뮤직카우, 투자자보호위원회 꾸려. 한국경제신문. <https://www.hankyung.com/it/article/202204053844i>
- 허준영, 양진용. (2015). SVM 기반의 재무 정보를 이용한 주가 예측. 정보과학회 컴퓨터의 실제 논문지, 21(3), 167-172. <https://doi.org/10.5626/KTCP.2015.21.3.167>

[국외 문헌]

- A Team. (2021, November 22). Artworks in Bits and Pieces:
- The Growing Trend of Fractional Ownership of Art in Korea. <https://k-artnow.com/artworks-in-bits-and-pieces-the-growing-trend-of-fractional-ownership-of-art-in-korea/>
- Aggarwal, A., Gupta, I., Garg, N., & Goel, A. (2019, August). Deep learning approach to determine the impact of socio economic factors on bitcoin price prediction. In 2019 twelfth international conference on contemporary computing (IC3) (pp. 1-5). IEEE.
- Art Basel. (2021, March 16). The Art Basel and UBS Global Art Market Report. <https://www.artbasel.com/stories/art-market-report-2021>
- Devadoss, A. V., & Ligor, T. A. A. (2013). Forecasting of stock prices using multi layer perceptron. International Journal of Computing Algorithm, 2, 440-449.
- Huck, N. (2009). Pairs selection and outranking: An application to the S&P 100 index. European Journal of Operational Research, 196(2), 819 - 825.

- Huck, N. (2010). Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, 207(3), 1702-1716.
- Kim, Y. (2021, March 29). Not just window shopping: Young generation dives into art market. *The Korea Herald*. <https://www.koreaherald.com/view.php?ud=20210329000839>
- Lowies, B., Whit, R. B., Viljoen, C., & McGreal, S. (2018). Fractional ownership - an alternative residential property investment vehicle. *Journal of Property Investment & Finance*. 36(6). 513-522. <https://doi.org/10.1108/JPIF-02-2018-0013>
- McNally, S., Roche, J., & Caton, S. (2018, March). Predicting the price of bitcoin using machine learning. In 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP) (pp. 339-343). IEEE.

Abstract

Price Prediction of Fractional Investment Products Using LSTM Algorithm: Focusing on Musicow

Hyunjo Jung* · Jaehwan Lee** · Jihae Suh***

Real estate and artworks were considered challenging investment targets for individual investors because of their relatively high average transaction price despite their long investment history. Recently, the so-called fractional investment, generally known as investing in a share of the ownership right for real-life assets, etc., and most investors perceive that they actually own a piece (fraction) of the ownership right through their investments, is gaining popularity. Founded in 2016, Musicow started the first service that allows users to invest in copyright fees related to music distribution. Using the LSTM algorithm, one of the deep learning algorithms, this research predict the price of right to participate in copyright fees traded in Musicow. In addition to variables related to claims such as transfer price, transaction volume of claims, and copyright fees, comprehensive indicators indicating the market conditions for music copyright fees participation, exchange rates reflecting economic conditions, KTB interest rates, and Korea Composite Stock Index were also used as variables. As a result, it was confirmed that the LSTM algorithm accurately predicts the transaction price even in the case of fractional investment which has a relatively low transaction volume.

Key Words : LSTM, Fractional Investment, Price Prediction, Musicow, Prediction Model.

Received : September 2, 2022 Revised : October 4, 2022 Accepted : October 31, 2022

Corresponding Author : Jihae Suh

* College of Business Administration, Seoul National University of Science and Technology

** Institute of Management Research, Seoul National University

*** Corresponding Author: Jihae Suh

College of Business Administration, Seoul National University of Science and Technology

232 Gongneung-ro, Nowon-gu, Seoul 01811, Korea

Tel: +82-2-970-6442, Fax: +82-10-8770-3593, E-mail: jihae@seoultech.ac.kr

저 자 소개



정현조

현재 서울과학기술대학교 경영학과 학사과정에 재학 중이다. 관심분야는 빅데이터 분석, MIS, 금융공학, 투자론, 인공지능 등이다.



이재환

서울대학교 경영학과에서 박사학위를 취득하였다. 현재 서울대학교 경영연구소 객원연구원으로 있으며 주요 관심분야는 비즈니스 애널리틱스, 딥러닝, 오피니언 마이닝 등이다.



서지혜

현재 서울과학기술대학교 조교수로 재직 중이다. 서울대학교에서 경영정보시스템 전공으로 박사학위를 취득하였다고 관심분야는 빅데이터 분석, 데이터 모델링 등이다. Decision Support Systems, Journal of Database Management, Electronic Commerce Research Applications, The Data Base for Advances in Information Systems 등 국내외 저널에 다수의 논문을 게재하였다.