

# Land NFTs Analysis & Forecasting

AI 18기 김보미

# Our **SERVICE**



01  
Problem  
Situation



02  
EDA &  
Vlsualization



03  
Feature  
Engineering



04  
Modeling



05  
Conclusion

- randomforest
- LSTM
- (Weighted Average)  
randomforest + LSTM

# Problem Situation

최근 NFT 시장에서는 급격한 가격 상승과 함께 많은 관심을 받았지만,  
이에 대한 비판과 함께 가격의 급락도 이어졌습니다.

저는 NFT 시장의 일시적인 변동성보다는 NFT가 가지는 중장기적인 기술적 혁신과  
활용 가능성에 더 관심이 있습니다

이에 따라서, 제가 작년에 구매한 토지 NFT인 소량주의 Orbcity의 가격 예측을 위해  
토지 NFT 대량주인 Decentraland와 Sandbox의 대량주 데이터를 활용하여  
land NFT의 분석과 더불어 전반적인 시장 동향을 파악하고자합니다.

또한 변동성이 매우 큰 NFT 시장에서 단기적인 투자 전략 수립을 위해,  
1일, 5일 뒤의 종가 예측도 함께 시도하려 합니다.

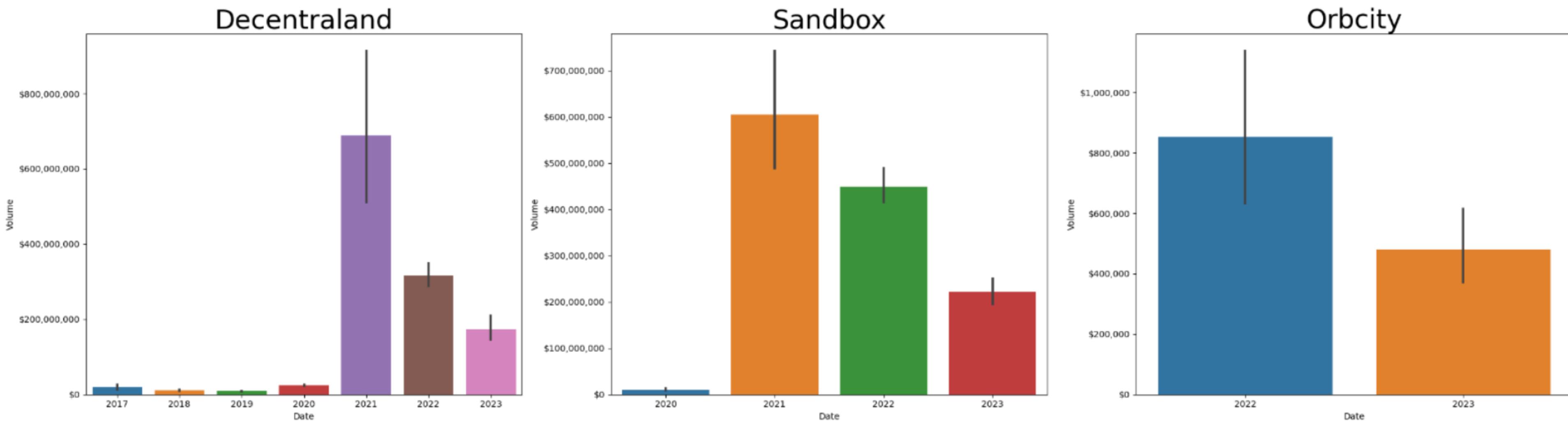
# EDA - Analysis



비슷한 시점에 급등, 급락.  
(2021년말) -> 이유는?

- 우크라이나 러시아 전쟁으로 코인 가격 하락 -> NFT 가치 하락
- 러그풀로 신뢰도 하락
- NFT 법률 지원 미비

# Dataset : 3 Land NFTs



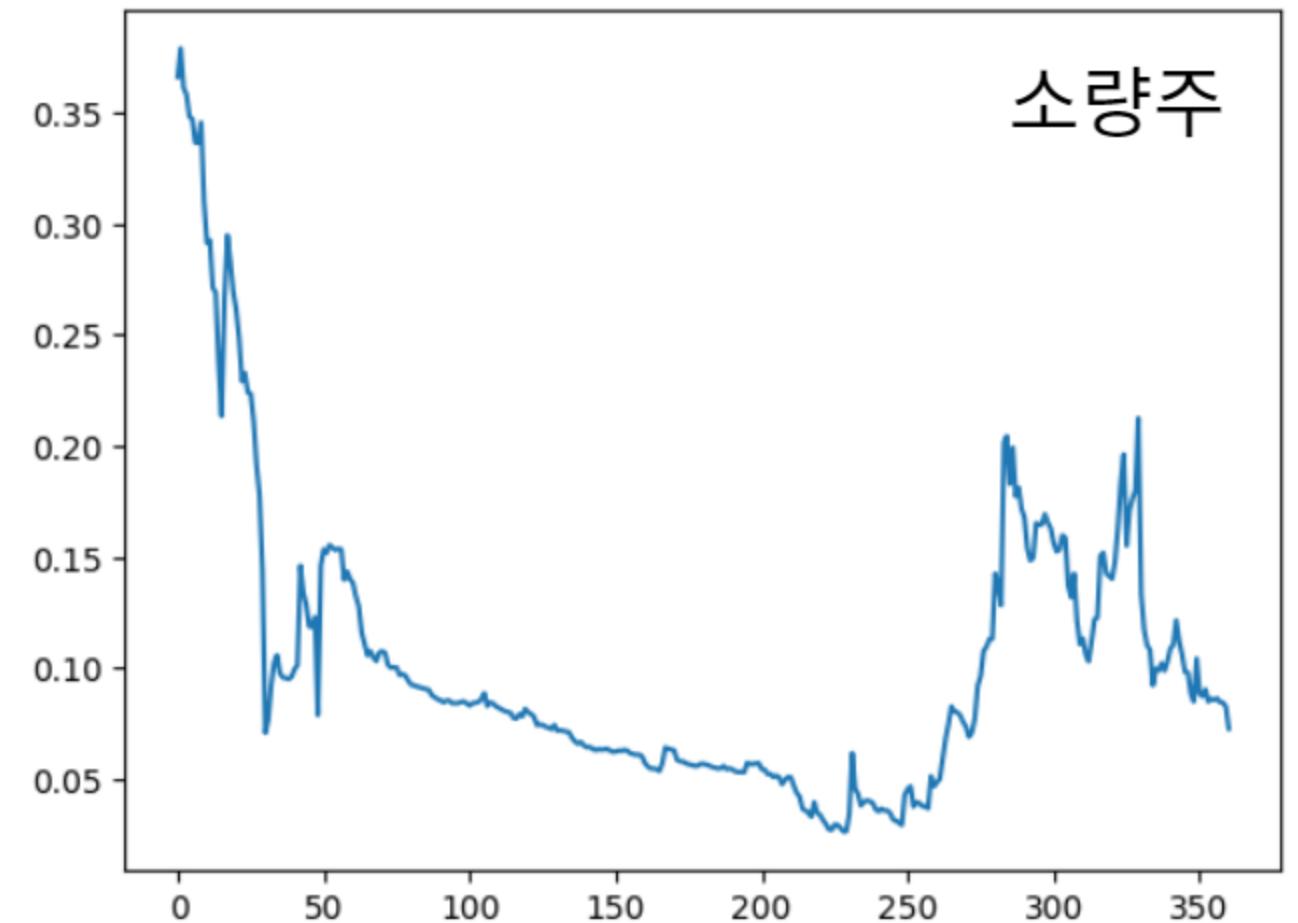
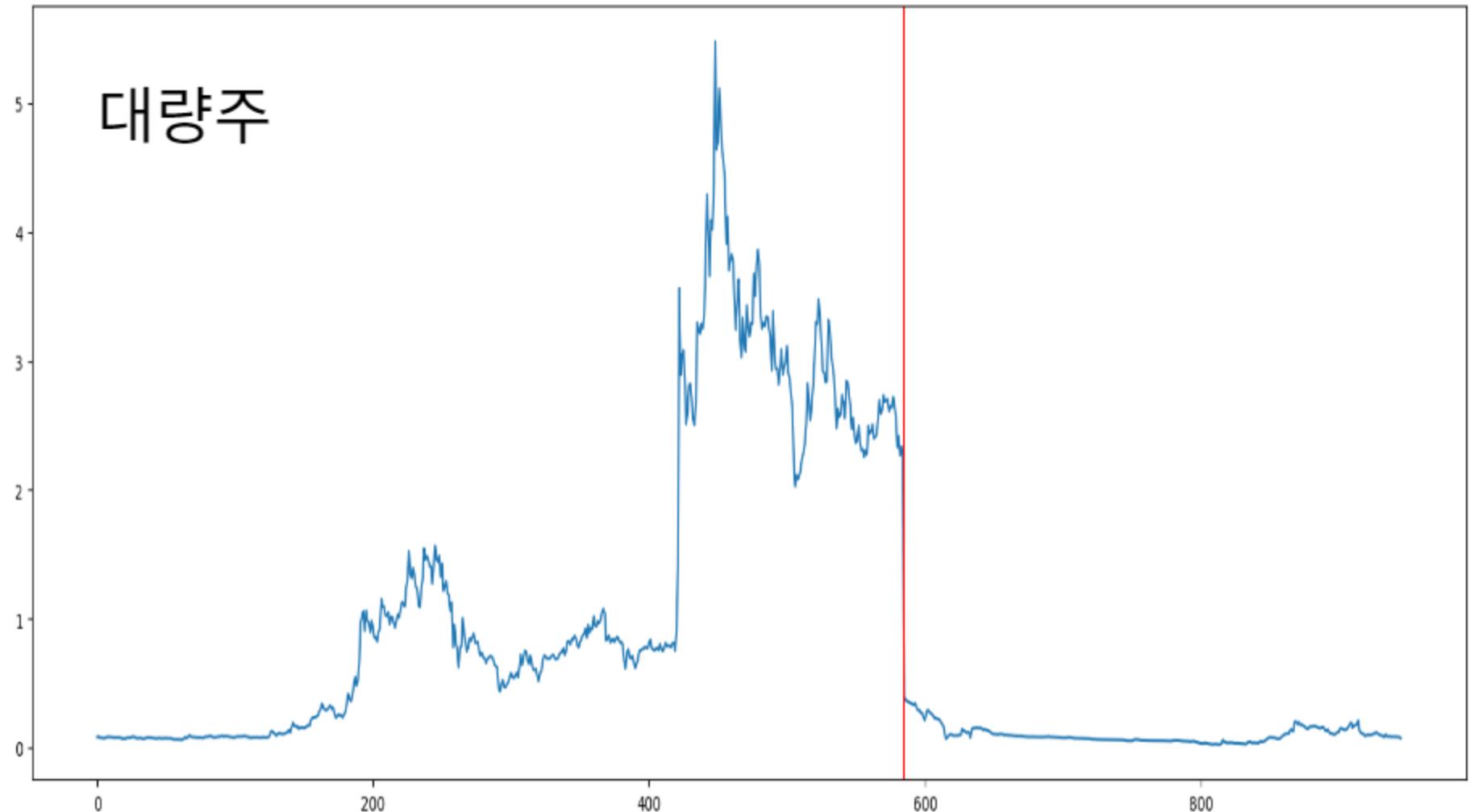
- 가정 : Land NFT 간의 유사성이 높다.

대량주 데이터(Decentraland, Sandbox) 이용해 소량주 데이터(Orbcity) 예측

-> 전이학습 : 대량주 데이터를 이용해 더 많은 정보를 얻고, 이를 바탕으로 소량주 데이터 예측

=> 소량주 데이터의 한계 극복, 더욱 정확한 예측 모델 구축

# Dataset : 3 Land NFTs



- 가정 : Land NFT 간의 유사성이 높다.

대량주 데이터(Decentraland, Sandbox) 이용해 소량주 데이터(Orbcity) 예측

-> 전이학습 : 대량주 데이터를 이용해 더 많은 정보를 얻고, 이를 바탕으로 소량주 데이터 예측

=> 소량주 데이터의 한계 극복, 더욱 정확한 예측 모델 구축

# EDA - Analysis

50-day data around the most active period of Sandbox



# Dataset - 소량주 Orbcity

Orbcity



# EDA

- - decentraland : (946,19)
- - sandbox : (946,19)
- - orbcity : (946,19)

대량주 : decentraland, sandbox

소량주 : orbcity

decentraland -> sandbox의 시작년도와 맞추어주었다.

	Decentraland	Sandbox	Orbcity
흥행 시작 년도	2017 말	2020 중	2022 초
데이터 개수	1975	966	382
최근 마지막 거래량	129775680	133827632	430847
거래 토큰	MAMA, LAND	SAND	ORB
스마트 계약 플랫폼	Ethereum	Ethereum	Binance
VR 지원	지원	지원	미지원
최근 대략 시가총액	\$1,110,824,414	\$134,447,000	\$9,358,000
총 시가 순위	대략45위	대략 55위	x
트위터 팔로워	631,868	1,095,543	45,976

# Feature Engineering

- Target Column
  - target\_1d : 다음날 종가 예측
  - target\_5d : 5일 뒤 종가 예측
- 보조 지표
  - 이동평균선 (SMA) (5,20) : 최근 \_일 간의 종가 평균치를 이용한 표
  - 지수이동평균선 (EMA) (12,26) : 최근 \_일 간의 가격 데이터를 가중치 부여해 계산한 평균치.
  - 이동평균수렴확산 (MACD) : 추세 전환 예측 지표
  - Signal : 추세 전환 신호
  - 상대강도지수 (RSI) : 가격 움직임의 상승, 하락 비교-> 과매수, 과매도 상태 파악 지표
  - Bollinger Bands : 주가의 변동성 이용-> 상한선, 하한선 그림 -> 과매수, 과매도 상태 파악 지표
  - cross : 이동평균선 간의 교차점을 나타내는 지표 (0이면 하락추세 , 1이면 상승추세)

# Modeling - Decentraland, Sandbox

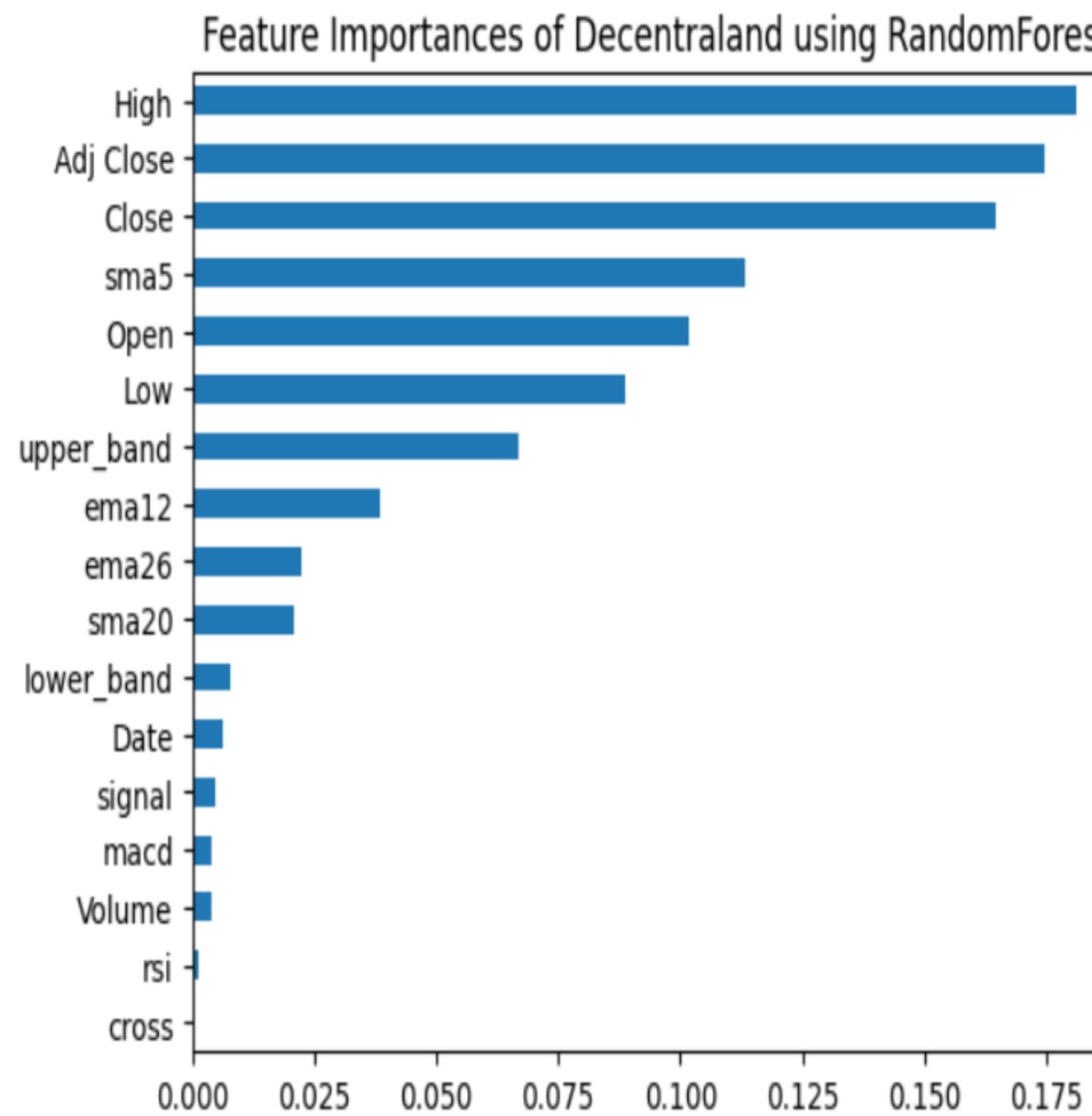
2년치 과거 데이터 -> 7개월 테스트 데이터 예측

train set : 2020-09-02 ~ 2022-09-27 (756개)  
test set : 2022-09-28~ 2023-04-05 (190개)

정규화 : MinMaxScaler

기준 모델 : 평균 사용 (모델 예측의 하한선)  
R2 of base model 0.0  
MAE of base model 0.85  
RMSE of bse model 1.08

# Modeling 1 - RandomForestRegressor



Decentraland dataset

- 1 일후 종가 예측
  - GridsearchCV
  - Feature importance

RandomForestRegressor(  
criterion='squared\_error',  
max\_depth=8,  
max\_features='sqrt',  
min\_samples\_leaf=3,  
min\_samples\_split=3,  
n\_estimators=1000,  
oob\_score=True,  
bootstrap=True  
random\_state=42)

- 5일 후 종가는 예측 성능이 매우 안 좋았다.

# Modeling 1 - RandomForestRegressor

RMSE : 0.0623

R-squared: 0.8104

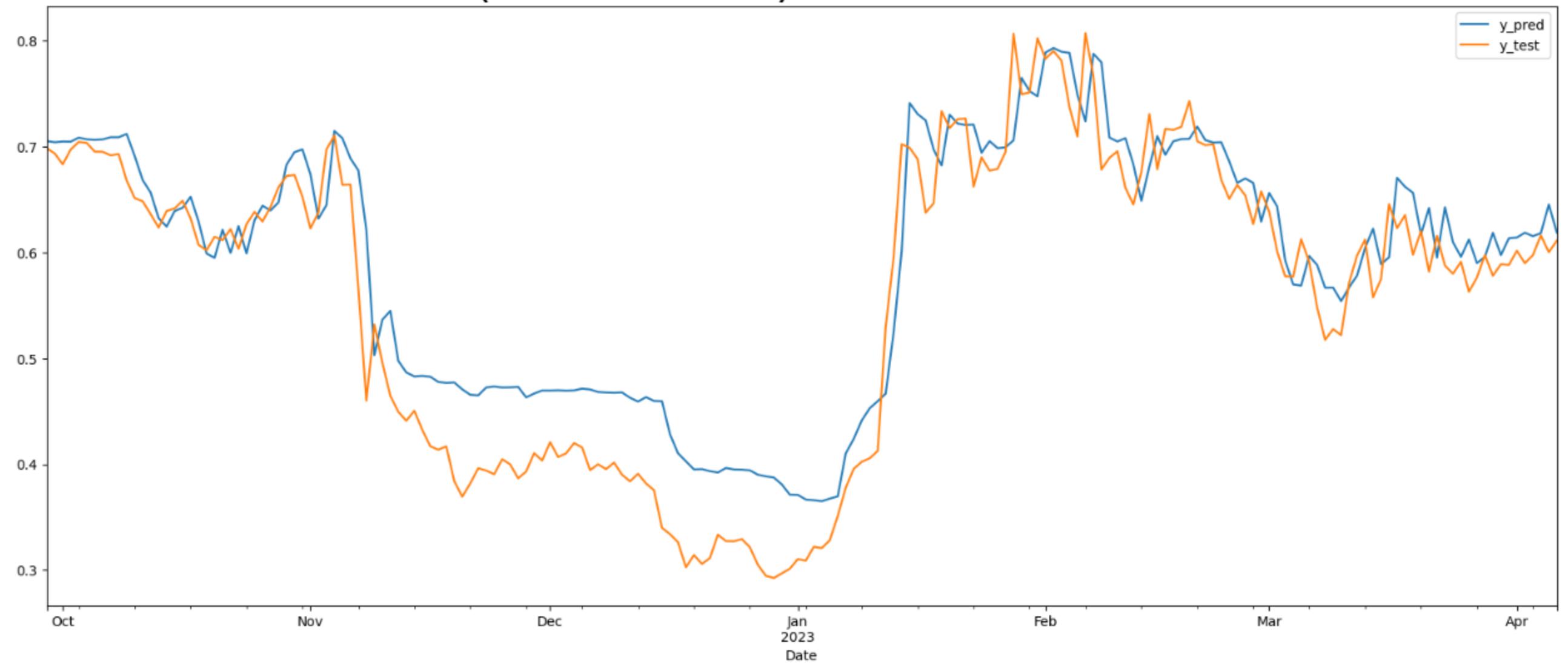
-> 중요치 않은 특성 drop 후,

RMSE : 0.0513

R-squared: 0.8713

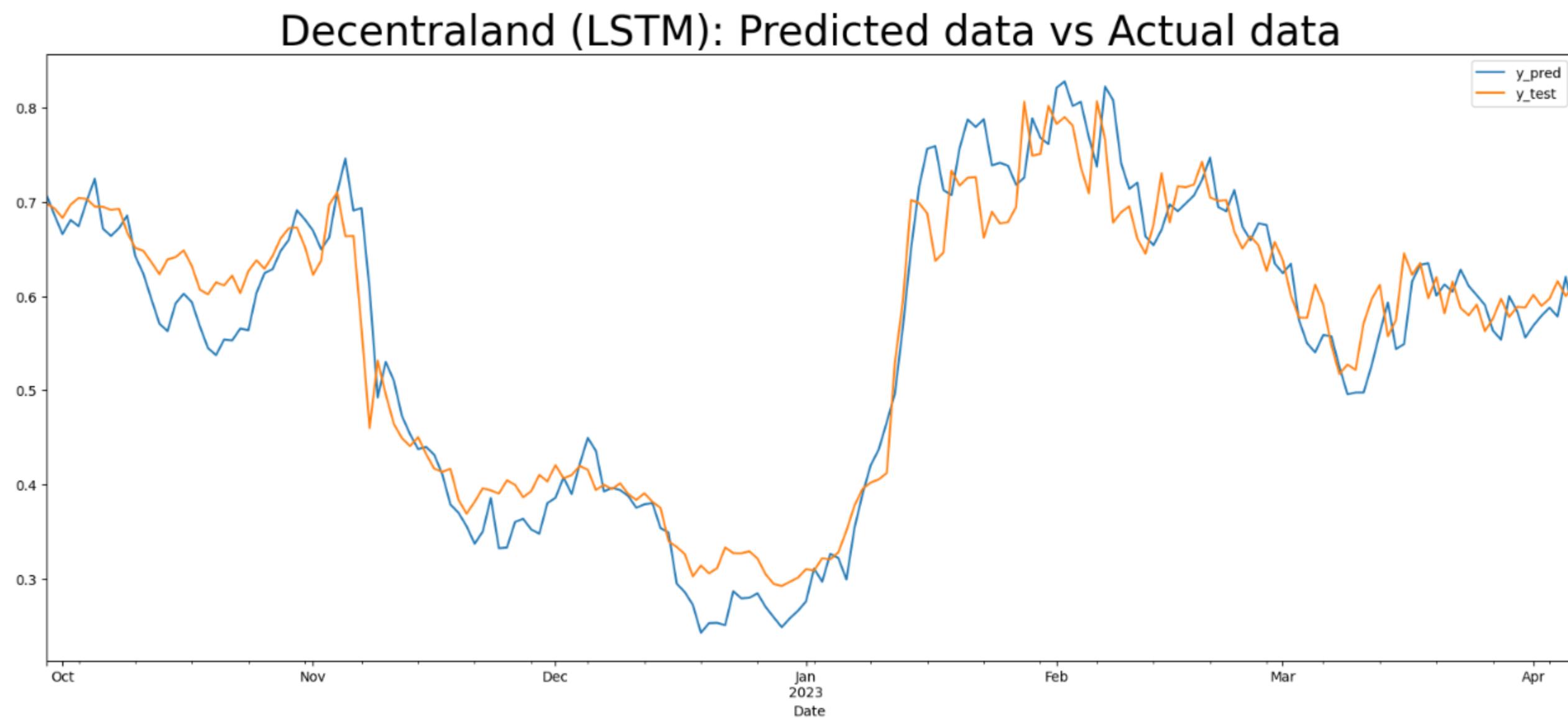
성능 향상됨.

Decentraland (Randomforest): Predicted data vs Actual data



# Modeling 2 - LSTM

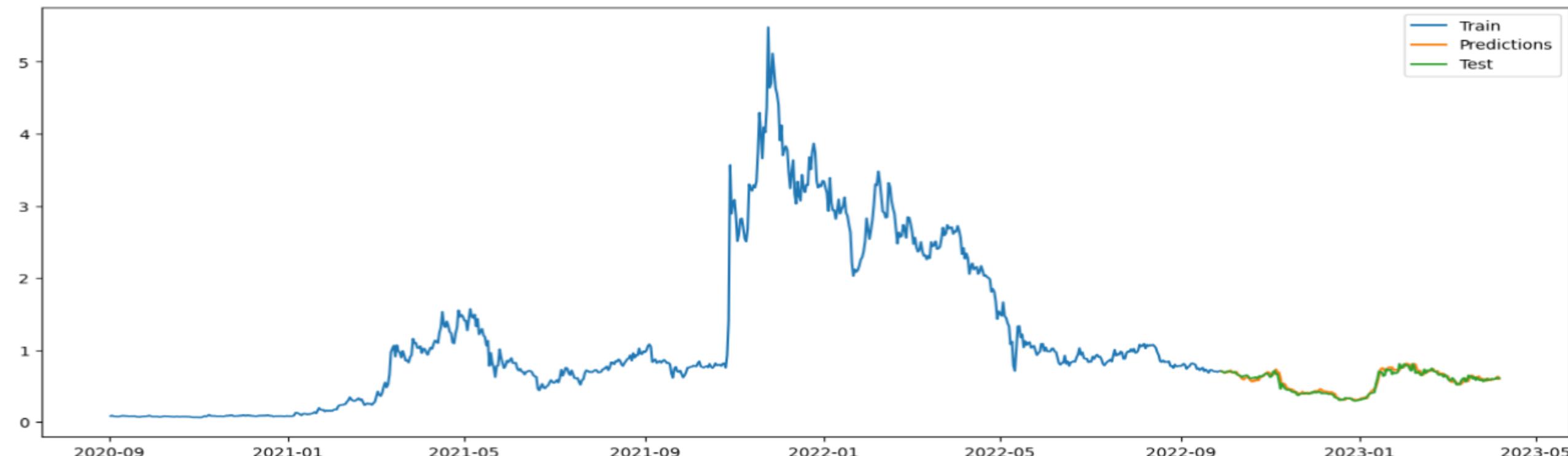
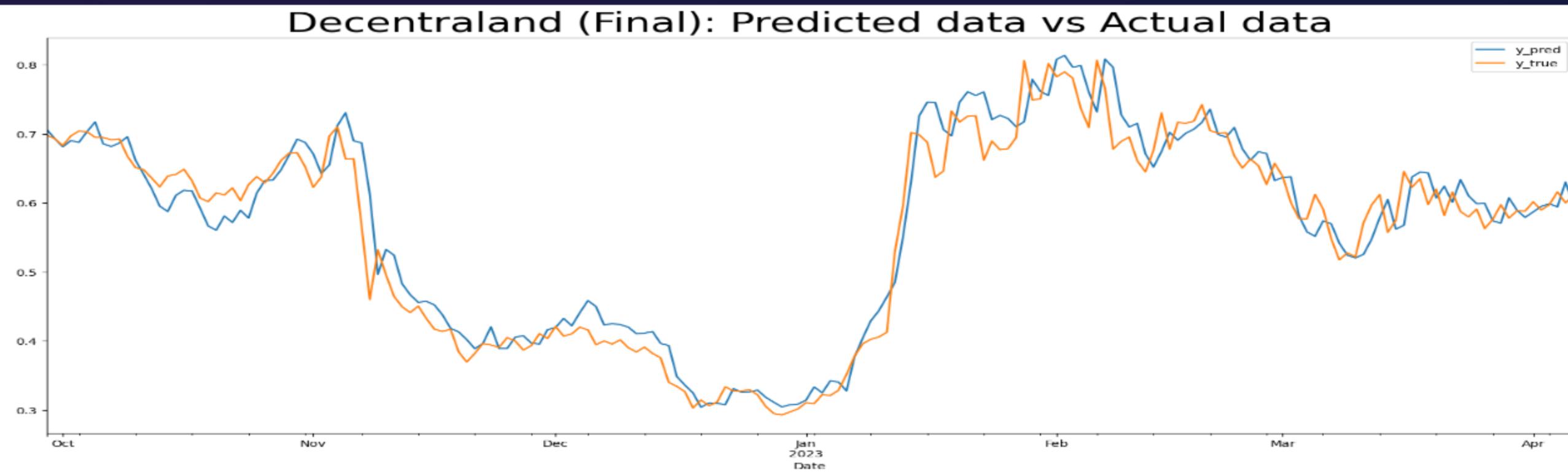
- 1일 뒤 종가 예측
    - RMSE : 0.0433
    - R-squared: 0.9082
  - 5일 뒤 종가 예측
    - 1일 뒤보다 성능이 좋지 않다.
    - RMSE : 0.0846
    - R-squared: 0.6449
- > randomforest보다 좋은 성능.



# Modeling 3 - (Weighted Average) RF+LSTM

가중치를 달리 주었더니,  
더 좋은 성능!!!!!!

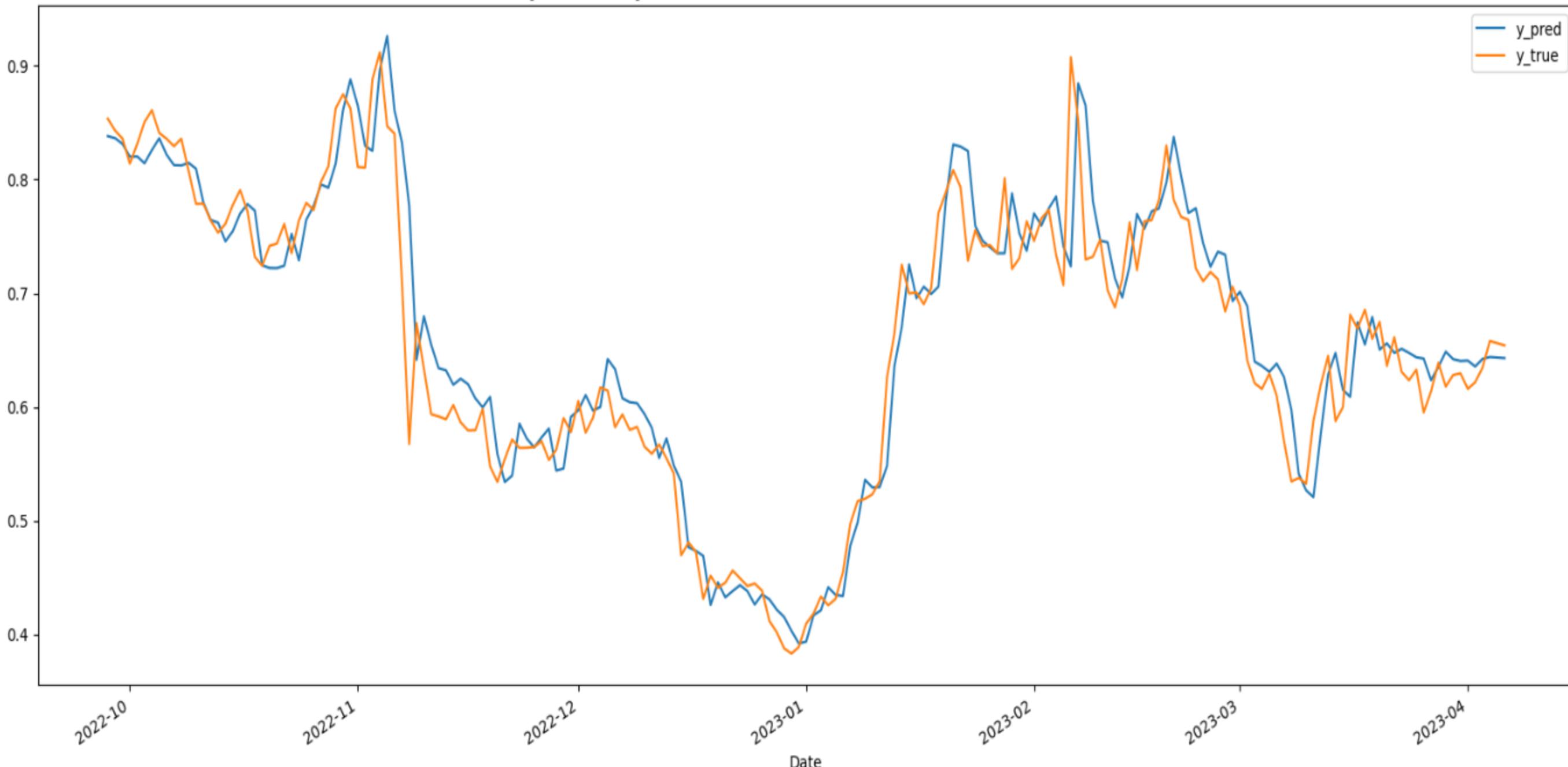
- lstm
  - RMSE : 0.0433
  - R-squared: 0.9082
- randomforest
  - RMSE : 0.0513
  - R-squared: 0.8713
- Weighted Average -  
randomforest(40%)+ lstm(60%)
  - RMSE: 0.0359
  - R-squared: 0.9369



# Modeling - Sandbox Dataset

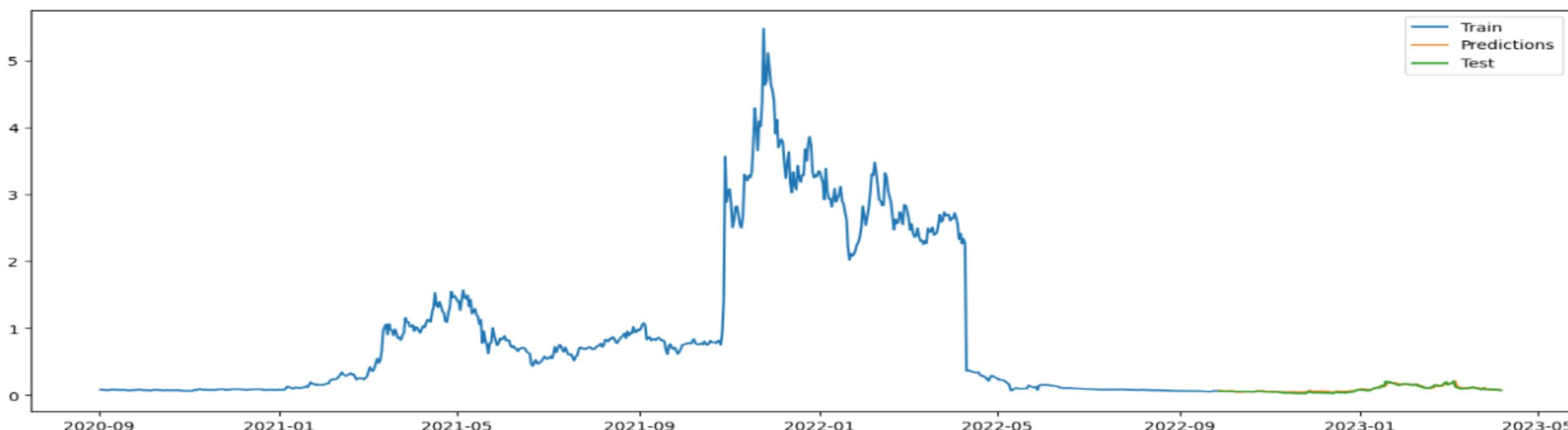
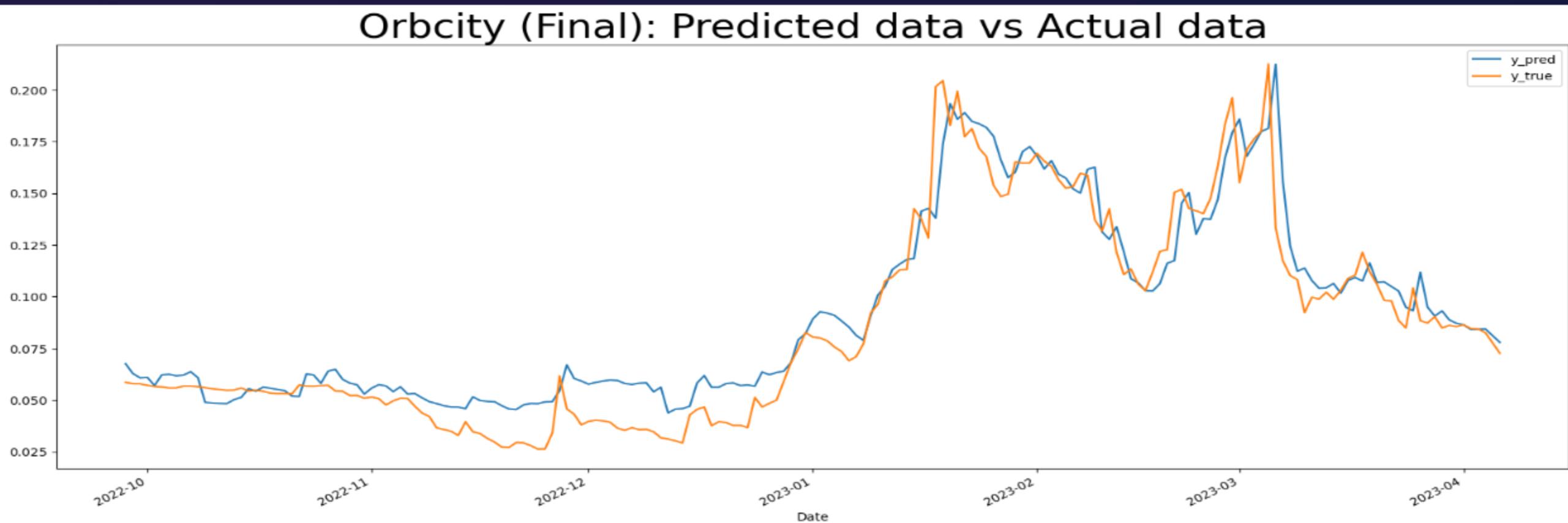
- LSTM
  - rmse: 0.0503
  - r2: 0.8427
- RandomForest
  - rmse: 0.0396
  - r2: 0.9025
- **LSTM(58%)+RF(42%)**
  - rmse: 0.0396
  - r2: **0.9074**

Sandbox (Final): Predicted data vs Actual data



# Modeling - Orbcity Dataset

- LSTM
  - rmse: 0.0216
  - r2: 0.8033
- RandomForest
  - rmse: 0.0160
  - r2: 0.8923
- **LSTM(58%)+RF(42%)**
  - rmse: 0.0147
  - r2: 0.9083



# Modeling Summary

	Decentraland	Sandbox	Orbcity	
	대량주	대량주	소량주	
RandomForest	0.8713	0.9025	0.8923	<ul style="list-style-type: none"><li>• RandomForest : 다수의 결정트리 이용해 예측을 수행해, 분산된 데이터에서도 좋은 성능을 보이며, 다양한 데이터 유형에서 잘 작동</li><li>• LSTM : 시계열 데이터에서 잘 작동하며, 과거 시계열 데이터를 이용해 예측을 수행하므로 데이터의 흐름을 파악하는 능력이 뛰어남.</li></ul>
LSTM	0.9082	0.8427	0.8033	
RF + LSTM	0.9369	0.9074	0.9083	각각의 모델의 장점을 활용해 가중 평균을 취하면, 모델의 성능이 더 좋아질 수 있다. 또한, 각 모델의 성능을 고려하여 가중치를 조절할 수 있어, 더욱 정확한 예측이 가능한다.

# Conclusion

- 가정 검증 완료
  - 대량주 데이터를 소량주 데이터의 과거에 적용하는 것이 더 나은 성능을 보인다는 가정이 검증되었습니다.
  - 이는 land nft 간의 유사성이 있다는 것을 의미하며, 이를 활용하여 소량주 데이터를 예측하는 데 효과적인 방법을 제시
- 효과
  - 대량주와 소량주 간의 관련성을 파악하여 비즈니스 전략을 개선하거나 새로운 제품 라인을 개발하는 등의 의사 결정
  - 재고를 최적화하고 재고 비용을 절감하며, 공급망을 최적화하여 제품 공급에 대한 효율성을 높일 수 있습니다.
  - 향후 NFT 시장의 변동성에 대한 인사이트를 제공
    - 가격은 이동평균선, 지수이동평균선, 볼린저밴드 등의 지표에 민감하게 반응할 수 있으며, 이러한 지표를 활용하여 투자 전략을 구성하는 것이 중요하다.
- 한계점
  - 해당 모델로는, 장기적인 미래 예측 힘들다.
  - 데이터의 부족과 모델의 결과가 예측과 가설일 뿐이며, 실제 투자에 대한 결정은 매우 복잡하고 개인적인 판단 필요
  - 해당 결과는 투자 전략의 참고 자료로 활용
  - NFT 커뮤니티, 이슈, 프로젝트, 게임 등 다른 요인에 대한 고려도 필수적입니다.
- 추가사항
  - 이후에, 장기적인 미래 예측과 원금 손실에 대한 예측 등 추가적인 분석을 토대로 더욱 정확하고 유용한 예측 모델을 개발하고 싶다.